



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Αγρονόμων και Τοπογράφων Μηχανικών
Τομέας Τοπογραφίας – Εργαστήριο Τηλεπισκόπησης

ΔΠΜΣ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ

Αυτόματη Ανίχνευση Εύφλεκτων Αερίων (C_3H_8 , CH_4) σε Θερμικά Υπέρυθρα Δεδομένα Βίντεο

Μεταπτυχιακή Εργασία

Αντώνιος Κουτρομπάς

Αθήνα, Οκτώβριος 2019



.....
Αντώνιος Κουτρομπάς

Πτυχιούχος Μηχανικός Παραγωγής και Διοίκησης

Copyright © Αντώνιος Κουτρομπάς, 2019.

Με επιφύλαξη παντός δικαιώματος, All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσοβίου Πολυτεχνείου



Μεταπτυχιακή Εργασία
ΔΠΜΣ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ

**Αυτόματη Ανίχνευση Εύφλεκτων Αερίων (C_3H_8 ,
 CH_4) σε Θερμικά Υπέρυθρα Δεδομένα Βίντεο**

Αντώνιος Κουτρομπάς
Α.Μ.: 60172215

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

.....
Δρ. Κ. Καραντζαλος
Αν. Καθηγητής, ΕΜΠ

.....
Δρ. Δ. Αργιαλάς
Καθηγητής, ΕΜΠ

.....
Δρ. Χ. Κοντοές
Διευθ. Ερευνών,
ΙΑΑΔΕΤ

.....
Αντώνιος Κουτρομπάς

Ημερομηνία Εξέτασης: Οκτώβριος 2019



MASTER THESIS
IN GEOINFORMATICS

**Automatic Detection of Flammable Gases (C_3H_8 ,
 CH_4) in Thermal Infrared Data Video**

Antonios Koutroumpas

R.N.: 60172215

EXAMINATION COMMITTEE

.....

Dr K. Karantzas
Assistant Professor, NTUA

.....

Dr D. Argialas
Professor, NTUA

.....

Dr. Kontoes
Research Director,
IAASARS

.....

Antonios Koutroumpas

Examination Date: October 2019



Περίληψη

Η χρήση θερμικών καμερών που λειτουργούν στο υπέρυθρο φάσμα, γίνεται ολοένα και πιο δημοφιλής για την ανάπτυξη τεχνικών απεικόνισης αερίων καθώς με την πάροδο του χρόνου αναδεικνύεται ως ένας γρήγορος, εύκολος και ασφαλής τρόπος απόκτησης δεδομένων για το σκοπό αυτό. Οι πληροφορίες που εξάγονται από το υπέρυθρο φάσμα διαδραματίζουν σημαντικό ρόλο σε πάρα πολλές εφαρμογές, όπως στην έγκαιρη ανίχνευση διαρροών σε πετρελαϊκές μονάδες ή σε υποσταθμούς μεταφοράς υπερυψηλής τάσης, συμβάλλοντας έτσι στην προστασία του περιβάλλοντος με την αποφυγή δυσμενών επιπτώσεων. Η πιο πρόσφατη γενιά θερμικών καμερών ανίχνευσης αερίων, χρησιμοποιεί ένα σύστημα υψηλής ευαισθησίας που στηρίζεται στην ανίχνευση της διαφοράς δύο θερμικών εικόνων υπέρυθρου φάσματος με την εφαρμογή δύο διαφορετικών φίλτρων απορρόφησης, έτσι ώστε τα αέρια να γίνονται περισσότερο ευκρινή. Παρόλη την πρωτοποριακή τεχνολογία του συστήματος, το οποίο έχει και υψηλό κόστος λόγω της πολυπλοκότητας της σύνθεσης των φίλτρων, οι εικόνες που παράγονται έχουν αρκετό θόρυβο, άρα είναι χαμηλής ποιότητας. Συνεπώς, τίθεται η αναγκαιότητα ανάπτυξης ενός συστήματος όπου από τις αρχικές θερμικές εικόνες θα είναι ικανό να επισημαίνει τη διάχυση των αερίων, παρέχοντας μια σαφή και καθαρή απεικόνισή τους. Στόχος της παρούσας διπλωματικής, είναι ο σχεδιασμός ενός τέτοιου συστήματος αυτόματης ανίχνευσης και απεικόνισης αερίων, μέσω της ανάπτυξης δύο επιμέρους μεθοδολογιών. Η πρώτη τεχνική βασίζεται σε μεθόδους υπολογιστικής μάθησης και υπολογιστικής όρασης, όπου τα κινούμενα αντικείμενα στα στιγμιότυπα ενός βίντεο εντοπίζονται και κατηγοριοποιούνται ως αέρια ή μη, με κριτήρια την υφή και τα χωροχρονικά χαρακτηριστικά τους. Η δεύτερη προσέγγιση στηρίζεται σε μεθόδους ψηφιακής επεξεργασίας εικόνες, όπου με την εξαγωγή του υποβάθρου της σκηνής από το συνδυασμό βίντεο καταγραφής ακτινοβολίας στο ορατό και στο υπέρυθρο φάσμα, πραγματοποιείται εντοπισμός του αερίου σε επίπεδο εικονοστοιχείου. Επιπρόσθετα, στην παρούσα διπλωματική παρουσιάζεται μία εκτενή ανάλυση των παραπάνω προτεινόμενων μεθοδολογιών, συγκρίνεται η απόδοση κατηγοριοποίησης χωροχρονικών περιγραφέων από δύο διαφορετικούς ταξινομητές και αξιολογείται η ικανότητα ανίχνευσης αερίων στα στιγμιότυπα του βίντεο της μεθόδου εξαγωγής υποβάθρου. Οι προτεινόμενες τεχνικές επέφεραν ικανοποιητικά αποτελέσματα επιτυχημένων ανιχνεύσεων, συνοδευόμενες και από τις αντίστοιχες επιμέρους εσφαλμένες ανιχνεύσεις.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Μηχανική Μάθηση, Μηχανική Όραση, Ψηφιακή Επεξεργασία Εικόνες, Αυτόματη Ανίχνευση Αερίων, Χωροχρονικοί Περιγραφείς Χαρακτηριστικών, Υπέρυθρο φάσμα

Abstract

The use of thermal cameras operating in the infrared spectrum is becoming more and more popular over time in the development of gas imaging techniques, as it is a fast, easy and secure way of remotely obtaining data for this purpose. The information extracted from the infrared play has a major role in a variety of applications, such as early detection of leaks at oil refineries or high voltage transmission at substations, thus protecting the environment subsequently avoiding adverse effects. The most recent generation of gas thermal cameras utilize a highly sensitive system based on detecting the differences between two infrared thermal images by applying simultaneously two distinct absorption filters in order to visualize the gas more clearly. Despite the innovation of this high cost technology, due to the complicated composition of filters, the result is insufficient because the produced images are full of noise resulting in low quality. Subsequently, an advanced system is needed to detect precisely the fuse of gases from the raw thermal images. The objective of this thesis is to design such an automatic gas detection and visualization system by suggesting two individual approaches. The first proposes a machine learning and computer vision technique, where moving objects in video frames are detected and classified as gases or non-gases, based on texture and spatiotemporal characteristics. As regards the second approach, which is based on digital imaging techniques, the background of the scene is extracted with background subtraction methods, combining synchronized video frames from visible and infrared spectrum, detecting gas at pixel level. Moreover, this thesis presents a wide analysis of the proposed methodologies, compares the categorization performance of spatiotemporal descriptors from two different classifiers and evaluates the gas detection efficiency of the background subtraction system. The evaluated methods of this thesis provide satisfactory gas detecting results followed by the corresponding individual false detections.

KEYWORDS: Machine Learning, Computer vision, Digital Imaging Techniques, Gas Detection, Infrared Spectrum, Spatiotemporal Feature Descriptors,

Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω θερμά όλους τους καθηγητές μου που με καθοδήγησαν και με κατεύθυναν κατά τη διάρκεια του μεταπτυχιακού προγράμματος και ιδιαίτερα τον επιβλέποντα αυτής της εργασίας καθηγητή Δρ. Κωνσταντίνο Καράντζαλο, για την εμπιστοσύνη που έδειξε στο πρόσωπό μου για την εκπόνηση της διπλωματική μου εργασίας καθώς και για την παρουσίαση ενός νέου ερευνητικού δρόμου, γεμάτου δυνατότητες.

Επίσης θα ήθελα να ευχαριστήσω τον διδάκτορα κ. Ζαχαρία Κανδυλάκη για την στήριξη και τη βοήθεια που μου παρείχε καθ' όλη τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας.

Θα ήθελα επιπλέον να εκφράσω τις ευχαριστίες μου στις εταιρείες Sierra Olympic και Revival για την παραχώρηση του εξοπλισμού και της θερμικής κάμερας για τη λήψη των βίντεο δεδομένων, καθώς και για την διάθεση των φιαλών των εξεταζόμενων αερίων, αντίστοιχα.

Τέλος, ένα μεγάλο ευχαριστώ στην οικογένεια μου για όσα μου έχουν προσφέρει καθώς και στην οικογένεια Ντουρτόγλου για την πολύτιμη συνεισφορά της που μετέτρεψε την πορεία των μεταπτυχιακών σπουδών μου σε ένα όμορφο ταξίδι.

Περιεχόμενα

Περίληψη.....	ii
Abstract	iii
Ευχαριστίες.....	iv
Περιεχόμενα	v
1 Πρόλογος	1
1.1 Εισαγωγή.....	1
1.2 Κίνητρο και Συνεισφορά.....	3
1.3 Δομή Διπλωματικής.....	5
2 Υπέρυθρη θερμική απεικόνιση	6
2.1 Υπέρυθρη Ακτινοβολία.....	6
2.2 Θερμικές Κάμερες	10
2.3 Θερμογραφία Υπέρυθρου στα Αέρια.....	12
3 Ανίχνευση στα στιγμιότυπα των βίντεο.....	16
3.1 Προεπεξεργασία εικόνας και αφαίρεση υποβάθρου	16
3.1.1 Αφαίρεση υποβάθρου μέσω τεχνικής διαδοχικών στιγμιότυπων	16
3.1.2 Αφαίρεση υποβάθρου μέσω στατιστικών μεθόδων	17
3.1.3 Αφαίρεση υποβάθρου μέσω επιλεκτικά μεταβαλλόμενου μοντέλου	17
3.2 Εξαγωγή Χαρακτηριστικών	18
3.3 Στατικοί Περιγραφείς Χαρακτηριστικών	19
3.3.1 Περιγραφέας Τοπικών Δυναδικών Προτύπων - Local Binary Patterns	19
3.3.2 Περιγραφέας Κυματιδίων - Wavelet	22

3.3.3	Περιγραφέας Ιστογράμματος Προσανατολισμού Ακμών - Edge Orientation Histogram.....	23
3.3.4	Περιγραφέας Ιστογράμματος Κλίσεων Διαβαθμίσεων - Histogram of Oriented Features	24
3.4	Δυναμικοί Περιγραφείς Χαρακτηριστικών	25
3.4.1	Περιγραφέας Ιστογράμματος Οπτικής Ροής - Histogram of Optical Flow	25
3.4.2	Περιγραφέας Τοπικών Δυαδικών Προτύπων Τριών Διαστάσεων - Local Binary Patterns from Three Orthogonal Planes	27
3.5	Αλγόριθμοι ταξινόμησης.....	28
3.5.1	Μηχανές Διανυσμάτων Υποστήριξης - Support Vector Machines.....	30
3.5.2	Μέθοδος AdaBoost	35
3.6	Αλγόριθμοι εξαγωγής και αντιστοίχισης χαρακτηριστικών	36
3.6.1	Αλγόριθμος SIFT	37
3.6.2	Αλγόριθμος RANSAC	37
3.7	Ανάλυση Συνδεδεμένων Στοιχείων	38
3.8	Μορφολογικές διαδικασίες Επεξεργασίας εικόνας	38
3.9	Βιβλιογραφική Ανάλυση: Ανίχνευση αντικειμένων μεταβαλλόμενου σχήματος ...	39
3.9.1	Σύνοψη συστημάτων θερμικής απεικόνισης αερίων.....	39
3.9.2	Σύνοψη συστημάτων ανίχνευσης καπνού σε βίντεο	40
4	Μεθοδολογία και Δεδομένα.....	43
4.1	Εξοπλισμός και απόκτηση δεδομένων	43
4.2	Τεχνική ανίχνευσης αερίου με χρήση ταξινομητών	45
4.2.1	Δημιουργία δεδομένων εκπαίδευσης	46
4.2.2	Προεπεξεργασία δεδομένων	47
4.2.3	Περιγραφείς Χαρακτηριστικών	47
4.2.4	Ταξινόμηση των περιοχών αερίου	51

4.2.5	Μετά-Επεξεργασία Αποτελεσμάτων.....	52
4.3	Τεχνική ανίχνευσης αερίου με αφαίρεση υποβάθρου.....	52
4.3.1	Συγχρονισμός των βίντεο	53
4.3.2	Ανίχνευση αερίου με εξαγωγή υποβάθρου.....	54
4.4	Αξιολόγηση Απόδοσης με χρήση Πίνακα Σύγκυσης.....	57
5	Αποτελέσματα.....	59
5.1	Μεθοδολογία αξιολόγησης αποτελεσμάτων	59
5.2	Δημιουργία δεδομένων επικύρωσης.....	60
5.3	Αποτελέσματα ανίχνευσης αερίου με χρήση ταξινομητών.....	61
5.4	Αποτελέσματα απόδοσης ανίχνευσης αερίου με εξαγωγή υποβάθρου.....	66
6	Συμπεράσματα	69
6.1	Ειδικά και Γενικά Συμπεράσματα	69
6.2	Προτάσεις για την εξέλιξη της παρούσας διπλωματικής	71
6.2.1	Διερεύνηση ποικιλίας του Συνόλου Δεδομένων	71
6.2.2	Βελτιώσεις στην εξαγωγή χαρακτηριστικών	72
6.2.3	Διαδικασία προεπεξεργασίας.....	72
6.2.4	Κινούμενη θερμική κάμερα.....	73
6.2.5	Βελτιώσεις στη μέθοδο εξαγωγής υποβάθρου.....	73
	Παράρτημα Α	74
	Βιβλιογραφία	79
	Λίστα Εικόνων.....	83
	Λίστα Πινάκων	87

Κεφάλαιο 1^ο

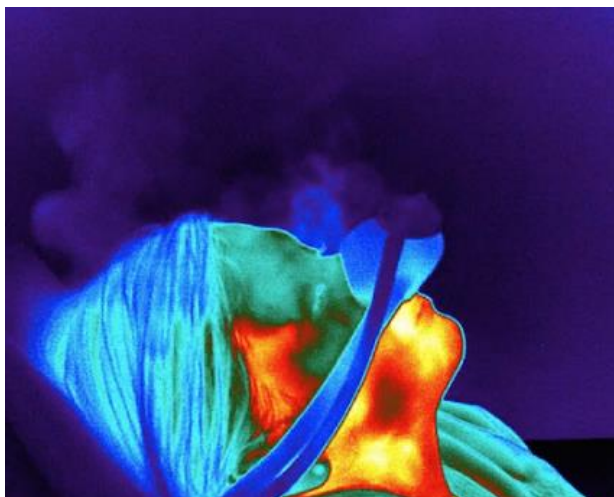
Στο κεφάλαιο αυτό γίνεται μια ιστορική αναδρομή σχετικά με την εξέλιξη της τεχνολογίας της θερμικής απεικόνισης, των πεδίων εφαρμογών της και ιδιαίτερα στον τομέα της ανίχνευσης αερίων, ενώ τέλος γίνεται μια συνοπτική αναφορά στη δομή της παρούσας διπλωματικής.

1 Πρόλογος

1.1 Εισαγωγή

Όταν παρατηρούμε το φάσμα της ηλιακής ακτινοβολίας που φθάνει στην επιφάνεια της γης, γίνεται φανερό ότι η ικανότητα ανίχνευσης του ανθρώπινου ματιού περιορίζεται σε ένα σχετικά στενό εύρος ζώνης του ηλιακού φάσματος, ή αλλιώς στο ορατό φάσμα, το οποίο κυμαίνεται μεταξύ των μηκών κύματος 0,38 μm και 0,78 μm . Διάφορα όμως ζώα, έχουν την ικανότητα να αντιλαμβάνονται μήκη κύματος που βρίσκονται εκτός του ορατού φάσματος του ανθρώπου και συνήθως στο υπέρυθρο τμήμα αυτού [1]. Λαμβάνοντας υπόψη τις αναλύσεις για τα μήκη κύματος εντός του οπτικού φάσματος, καθώς και το γεγονός ότι το μέγιστο της ακτινοβολίας βρίσκεται ακριβώς στο κέντρο του οπτικού φάσματος (σε μήκος κύματος 0,55 μm , το οποίο αντιστοιχεί στην αντίληψη μας για το πράσινο χρώμα), οι φασματικοί περιορισμοί του οπτικού συστήματος έχουν νόημα. Ωστόσο η υπέρυθρη ζώνη παρέχει πολύτιμες πληροφορίες που μπορούν να γίνουν εκμεταλλεύσιμες με τον κατάλληλο εξοπλισμό, όπως με τη χρήση θερμικών καμερών που απεικονίζουν την ηλεκτρομαγνητική ακτινοβολία που εκπέμπεται από όλα τα σώματα. Οι θερμικές κάμερες εισήχθησαν στην αγορά για πρώτη φορά το 1964, ενώ στο πέρασμα του χρόνου γίνονται όλο και πιο δημοφιλείς σε πολλούς διαφορετικούς τομείς. Η μέτρηση της θερμοκρασίας από απόσταση έχει αναπτυχθεί σε πολλά διαφορετικά πεδία, καλύπτοντας ένα ευρύ φάσμα στρατιωτικών, ιατρικών και άλλου κοινωνικού ενδιαφέροντος εφαρμογών. Πολλά παραδείγματα χρήσης θερμικών καμερών αφορούν στον εντοπισμό ατόμων μέσα από καπνό, διάγνωσης καρκίνου του μαστού [2], την αξιολόγηση της ενεργειακής αποτελεσματικότητας κτιρίων καθώς και τον προσδιορισμό της ύπαρξης και της θέσης διαρροών αερίων σε εργοστασιακές μονάδες. Επιπρόσθετα, η ανάπτυξη της γεωργίας ακριβείας καθώς και νέων τεχνολογιών, όπως τα μη επανδρωμένα αεροσκάφη, άνοιξε το δρόμο για τη χρήση θερμικών καμερών και στο πεδίο της γεωργίας. Στον τομέα αυτό η θερμική απεικόνιση έχει υιοθετηθεί με επιτυχία στη μελέτη της φυσιολογίας των φυτών, του προγραμματισμού της άρδευσης καθώς και στην πρόληψη εξάπλωσης ασθενειών. Από την επεξεργασία και την ανάλυση των δεδομένων που συλλέγονται από μια θερμική κάμερα και αφορούν στην υπέρυθρη ακτινοβολία που αντανακλά στα φυτά, μπορούν να αντιμετωπιστούν πολλά προβλήματα τα οποία προκαλούνται από τις ακραίες αλλαγές του κλίματος, την επέκταση ζιζανίων και παρασίτων καθώς και την ελλιπή εφαρμογή λιπασμάτων. Καθώς τα περισσότερα αέρια είναι αόρατα με γυμνό μάτι και συχνά άοσμα, οι διαρροές τους παρουσιάζονται είναι είτε άμεσου κίνδυνου, όπου απελευθερώνονται τοξικά ή αναισθητικά αέρια, είτε έμμεσου κίνδυνου, όπου απελευθερώνονται εύφλεκτα αέρια, με αποτέλεσμα να δημιουργούνται επισφαλείς συνθήκες. Σήμερα, υπάρχει μια μεγάλη ποικιλία θερμικών καμερών που αφορούν στην ανίχνευση διαφορετικών αερίων, πολλά από τα οποία ανήκουν στην κατηγορία των πτητικών οργανικών ενώσεων (VOCs). Οι πτητικές οργανικές ενώσεις είναι ενώσεις που έχουν υψηλή τάση ατμών σε συνηθισμένη θερμοκρασία δωματίου, ενώ περιλαμβάνουν φυσικές ή συνθετικές χημικές ενώσεις που σε πολλές περιπτώσεις είναι τοξικές

ή περιβαλλοντικά επιβλαβείς, όπως για παράδειγμα τα καύσιμα, οι διαλύτες ή τα ψυκτικά μέσα. Μερικά από τα πιο κοινά πτητικά αέρια στα οποία χρησιμοποιούνται θερμικές κάμερες για τον εντοπισμό τους είναι το μεθάνιο, το προπάνιο και το βενζόλιο, ενώ από τα μη πτητικά είναι το εξαφθοριούχο θείο, η αμμωνία, το μονοξείδιο και το διοξείδιο του άνθρακα. Συνεπώς, θερμικές κάμερες που αφορούν στον εντοπισμό αερίων έχουν ευρεία εφαρμογή στην πετρελαϊκή βιομηχανία, όπου και χρησιμοποιούνται για την επιθεώρηση αγωγών, σωληνώσεων και την ασφάλεια των εγκαταστάσεων παραγωγής. Ένα άλλο μεγάλο πεδίο εφαρμογής των θερμικών καμερών αφορά στην παρακολούθηση των υποσταθμών υπερυψηλής τάσης, που σχετίζεται κυρίως με διαρροές αερίων στην ατμόσφαιρα, όπως είναι το εξαφθοριούχο θείο (SF_6). Το SF_6 είναι ένα ανθρωπογενές αέριο και χρησιμοποιείται λόγω της ικανότητας του να “σβήνει” ακαριαία τα ηλεκτρικά τόξα στα καλώδια υψηλής τάσης, με εφαρμογή στα καλώδια υψηλής τάσης σαν μονωτικό στους υποσταθμούς υποβιβασμού τάσης και διανομής ηλεκτρικού ρεύματος και στους διακόπτες υψηλής τάσης σαν διηλεκτρικό. Επιπλέον το SF_6 είναι το πιο δραστικό απ’ όλα τα ανθρωπογενή αέρια ως προς τη δυναμικότητα επιδείνωσης του γνωστού “φαινομένου του θερμοκηπίου”. Οι συνηθισμένοι τρόποι ανίχνευσης διαρροών, όπως είναι οι μετρητές υψηλής τάσης, είναι αρκετά ασαφείς για να εξασφαλίσουν την τήρηση όλων των απαραίτητων κανόνων ασφαλείας. Οι θερμικές κάμερες από την άλλη πλευρά, είναι υψηλής ευαισθησίας συστήματα που δίνουν τη δυνατότητα εντοπισμού ακόμη και των πιο μικρών διαρροών, ενώ η τεχνολογία της οπτικής απεικόνισης αερίων πρόσφατα ανακοινώθηκε από την Υπηρεσία Προστασίας Περιβάλλοντος των Ηνωμένων Πολιτειών (Environmental Protection Agency) ως το βέλτιστο σύστημα για τη μείωση εκπομπών.



Εικόνα 1.1: Θερμική απεικόνιση διαρροής αναισθητικού αερίου από μια μάσκα [3].

Άλλα πεδία εφαρμογής των θερμικών καμερών εντοπισμού αερίων είναι η διασφάλιση της ποιότητας στις γραμμές παραγωγής προϊόντων που περιέχουν ψυκτικά αέρια. Εκτός από εφαρμογές στην βιομηχανία, η συγκεκριμένη τεχνολογία βρίσκει εφαρμογή όπως προαναφέρθηκε και σε πεδία της ιατρικής και της δημόσιας ασφαλείας, καθώς κάμερες απεικόνισης αερίων μπορούν να χρησιμοποιηθούν για τον εντοπισμό διαρροών αζώτου στα δίκτυα διανομής των νοσοκομείων. Η ανίχνευση μιας τέτοιας διαρροής με ένα συμβατικό σημειακό αισθητήρα είναι μια χρονοβόρα διαδικασία και εμπεριέχει τον κίνδυνο υπολογισμού μετρήσεων με υψηλές αποκλίσεις, καθώς ο χειριστής πρέπει να ανιχνεύει με το χέρι κάθε πιθανή διαρροή. Αντιθέτως, ο εντοπισμός διαρροών με τη χρήση θερμικών καμερών επιτρέπει τον προσδιορισμό της θέσης της πηγής της διαρροής καθώς και την ποσοτικοποίηση του εκλούμενου αερίου. Επιπρόσθετα, είναι πιθανή η εμφάνιση διαρροής αερίων αναισθησίας μέσα σε μία χειρουργική αίθουσα, εξαιτίας των χαλαρών συνδέσεων ή ελαττωμάτων της ίδιας συσκευής αναισθησίας.

1.2 Κίνητρο και Συνεισφορά

Μέχρι πρότινος, από τους πιο διαδεδομένους τρόπους ανίχνευσης διαρροών αερίων ήταν η χρήση ανιχνευτών καταγραφής σημειακών μετρήσεων υπέρυθρης ακτινοβολίας. Το πρόβλημα που ενέχει η μεθοδολογία αυτή είναι ότι στηρίζεται στο γεγονός ότι πρέπει να βρίσκεται ο αισθητήρας κοντά στην πηγή της διαρροής προκειμένου να παράγει ικανοποιητικά αποτελέσματα ανίχνευσης. Ενώ υπάρχει μια μεγάλη ποικιλία μεθοδολογιών ανίχνευσης αερίων, όπως είναι τα υπερηχητικά συστήματα, τα συστήματα που στηρίζονται στην αντανάκλαση ή σε χημικές αντιδράσεις, καθένα έχει μια σειρά μειονεκτημάτων, είτε λόγω της χαμηλής ευαισθησίας τους, είτε λόγω της δυσμετακίνητης φύσης τους. Για αυτό το λόγο, χρησιμοποιείται συνήθως ένας συνδυασμός διαφορετικών μεθόδων για την εξαγωγή ικανοποιητικών αποτελεσμάτων ανίχνευσης διαρροών. Πιο αναλυτικά, οι τεχνικές ανίχνευσης ακουστικών διαρροών με χρήση υπερηχητικών συστημάτων, έχουν μελετηθεί από το 1930 και απαιτούν μια μόνιμη εγκατάσταση μονάδων ανίχνευσης σε διαφορετικά τμήματα κατά μήκος ενός αγωγού. Σε κάθε τμήμα υπάρχουν δύο ανιχνευτές, ένας μετρητής ροής και ένας αισθητήρας θερμοκρασίας. Από τα δεδομένα που συλλέγονται υπολογίζεται το ποσοστό της ογκομετρικής ροής, τη θερμοκρασία περιβάλλοντος και αερίου, η ταχύτητα αναπαραγωγής του ήχου καθώς και άλλες διαγνωστικές μετρήσεις σχετικά με τη θέση των ανιχνευτών, προκειμένου να συγκριθούν με εκείνες κατά την εισροή και εκροή του αερίου από τον αγωγό.

Επιπλέον, οι σημειακοί αισθητήρες υπέρυθρων χρησιμοποιούν την ακτινοβολία που διέρχεται από ένα γνωστό αέριο συγκεκριμένου όγκου, δηλαδή την ενέργεια που απορροφάται σε ορισμένα μήκη κύματος, αναλόγως με τις ιδιότητες του συγκεκριμένου αερίου. Η ενέργεια που συγκεντρώνεται στο συγκεκριμένο μήκος κύματος συγκρίνεται με την ενέργεια ενός μήκος κύματος εκτός της ζώνης απορρόφησης του αερίου προς ανίχνευση και η διαφορά τους είναι ανάλογη της συγκέντρωσης του υπάρχοντος αερίου. Αντίστοιχα, οι ηλεκτροχημικοί ανιχνευτές αερίων επιτρέπουν στα αέρια να διέρχονται μέσω μιας πορώδους μεμβράνης σε ένα ηλεκτρόδιο το οποίο είτε χημικά οξειδώνεται είτε μειώνεται, ενώ το ποσοστό οξείδωσης του αερίου στο ηλεκτρόδιο υποδεικνύει τη συγκέντρωση του αερίου. Από την άλλη πλευρά, η μεθοδολογία της θερμικής απεικόνισης του αερίου παρουσιάζει πολλά πλεονεκτήματα σε σύγκριση με τις προαναφερθείσες μεθόδους. Η χρήση θερμικών καμερών δίνει τη δυνατότητα εντοπισμού της πηγής ή των πηγών διαρροών αερίων σε μια ευρεία εξεταζόμενη περιοχή, καθώς επίσης είναι εφικτή η αξιολόγηση του μεγέθους, της κατεύθυνσης διάδοσης και της ταχύτητας του αερίου.

Ο εντοπισμός πραγματοποιείται σε ασφαλή απόσταση, χωρίς να υπάρχει ο κίνδυνος έκθεσης του χειριστή σε δυνητικά επιβλαβή αέρια ή σε επικίνδυνα συστατικά τους. Μειονεκτήματα της μεθοδολογίας είναι η ευαισθησία της σε μικρομετακινήσεις της κάμερας, καθώς και το γεγονός ότι η απεικόνιση των θερμικών διαφορών μιας σκηνής δεν αποτελεί από μόνο του ένα ολοκληρωμένο σύστημα ανίχνευσης αερίων. Ένα σύστημα ανίχνευσης πρέπει να επεξεργάζεται τα αποτελέσματα της ανιχνευθείσας θερμικής ακτινοβολίας με τέτοιο τρόπο ώστε το τελικό αποτέλεσμα να παρέχει στον χρήστη απευθείας την πηγή της διαρροής με ένα ποσοστό σφάλματος καθώς και ένα συνδυαστικό τρόπο οπτικοποίησης των αποτελεσμάτων με απεικόνιση στο ορατό φάσμα.



Εικόνα 1.2: Σημειακός αισθητήρας ανίχνευσης μεθανίου (CH_4) και προπανίου (C_3H_8) της εταιρείας EURO-GAS.

Ο συνδυαστικός τρόπος οπτικοποίησης αποτελεί επιθυμητό χαρακτηριστικό για τις επόμενες γενιές καμερών ανίχνευσης αερίων, αφού ένα τέτοιο σύστημα θα διευκολύνει σημαντικά την ανίχνευση διαρροών και θα καταστήσει πιο αποτελεσματική τη χρήση των θερμικών καμερών αερίων. Σκοπός της παρούσας εργασίας είναι ο προσδιορισμός του συνόλου των χαρακτηριστικών που θα επιτρέπουν την περιγραφή των χαρακτηριστικών ενός αερίου, η ανάπτυξη ενός συστήματος ανίχνευσης που θα βασίζεται στην μηχανική μάθηση με χρήση δεικτών και ταξινομητών, η εκπόνηση μιας νέας μεθοδολογίας συνδυασμού καμερών ορατού και υπέρυθρου φάσματος και η αξιολόγηση των αποτελεσμάτων των δύο μεθοδολογιών ως προς την απόδοσή τους.

Στη παρούσα εργασία η πειραματική διάταξη που χρησιμοποιήθηκε αφορούσε στατικές κάμερες, καθώς μια κινητική εφαρμογή θα εισήγαγε πρόσθετες προκλήσεις. Ένα τέτοιο σύστημα θα μπορούσε να αποτελεί πρόταση εξέλιξης της παρούσας διπλωματικής, καθώς ιστορικά τα συστήματα αναγνώρισης με βάση την μηχανική μάθηση, όπως για παράδειγμα τα συστήματα αναγνώρισης προσώπου ή σώματος, αναπτύχθηκαν ξεκινώντας από στατικές διατάξεις καμερών και σήμερα αποτελούν εργαλείο, με υψηλά ποσοστά απόδοσης με ανίχνευση σε πραγματικό χρόνο, πολλών ηλεκτρονικών συσκευών (π.χ. κινητών τηλεφώνων, φορητών υπολογιστών). Ενώ ένας αλγόριθμος ανίχνευσης αερίων σε μία στατική διάταξη κάμερας θεωρείται ως ένα στάδιο πρώιμης ανάπτυξης από ένα πιο εξελιγμένο σύστημα, όπως μιας κάμερας χειρός, εντούτοις μπορεί να χρησιμοποιηθεί αυτόνομα, για την επίβλεψη περιοχών υψηλού κινδύνου, όπως είναι τα πετρελαϊκά διυλιστήρια που είναι ευαίσθητα σε διαρροές αερίων. Οι ερευνητικές προσεγγίσεις για την ανάπτυξη αλγορίθμων και συστημάτων ανίχνευσης είναι σχετικά λίγες, ενώ οι περισσότερες σχετίζονται στις μεθόδους επεξεργασίας εικόνων και της υπολογιστικής όρασης.



Εικόνα 1.3: Ενισχυμένη θερμική απεικόνιση υψηλής ευαισθησίας με χρήση θερμικής κάμερας Ventus της εταιρείας Sierra Olympic.

Η παρούσα εργασία, στο πρώτο μέρος της, παρουσιάζει μια προσέγγιση που στηρίζεται σε τεχνικές της υπολογιστικής μάθησης και όρασης, χρησιμοποιώντας μια θερμική κάμερα, ενώ στο δεύτερο παρουσιάζεται ένα νέο σύστημα συνδυασμού συγχρονισμένων ανιχνεύσεων στα τμήματα του ορατού και υπέρυθρου φάσματος. Το σενάριο που εξετάζεται είναι η διάχυση αερίου και η επίδραση περιβαλλοντικών παραγόντων σε αυτό όπως ο άνεμος και άλλα κινούμενα αντικείμενα στη σκηνή. Τα αέρια που χρησιμοποιήθηκαν στην πειραματική διάταξη ήταν το μεθάνιο και το προπάνιο, ενώ η σκηνή εκτός από την ροή του αερίου περιείχε

διελεύσεις περαστικών, οχημάτων καθώς και τον χειριστή ρύθμισης της βαλβίδας ροής των αερίων.

1.3 Δομή Διπλωματικής

Ο στόχος των πρώτων δύο κεφαλαίων είναι η εισαγωγή και η περιγραφή του θεωρητικού υποβάθρου της θερμικής απεικόνισης και της υπολογιστικής μάθησης. Στο δεύτερο κεφάλαιο αναλύεται συνοπτικά ένα σύνολο όρων και εννοιών σχετικά με την υπέρυθη ακτινοβολία και των φυσικών νόμων που τη διέπουν, των χαρακτηριστικών των θερμικών καμερών, ενώ επίσης γίνεται μια επανεξέταση των υφιστάμενων μεθόδων αυτόματης ανίχνευσης και απεικόνισης των αερίων. Στη συνέχεια, στο τρίτο κεφάλαιο, παρουσιάζονται οι αρχές ανίχνευσης αντικειμένων σε στιγμιότυπα από βίντεο, ενώ αναλύονται εκτενώς τα εργαλεία της υπολογιστικής μάθησης όπως τα διάφορα είδη δεικτών περιγραφής, οι αλγόριθμοι ταξινόμησης και εξαγωγής-αντιστοίχισης χαρακτηριστικών καθώς και οι τεχνικές ψηφιακής επεξεργασίας εικόνας. Επιπρόσθετα γίνεται μια εκτενή αναφορά στα υπάρχοντα συστήματα ανίχνευσης αερίων και καπνού. Στο κεφάλαιο τέσσερα, παρουσιάζονται οι μεθοδολογίες της παρούσας εργασίας, συμπεριλαμβανομένων των χαρακτηριστικών του εξοπλισμού που χρησιμοποιήθηκε στην πειραματική διάταξη, τη μεθοδολογία που ακολουθήθηκε σε κάθε στάδιο των δύο προτεινόμενων τεχνικών ανίχνευσης καθώς και μία ανάλυση των εργαλείων αξιολόγησης των τεχνικών. Στο πέμπτο κεφάλαιο, αναλύεται ο τρόπος δημιουργίας δεδομένων επικύρωσης, γίνεται παρουσίαση και σχολιασμός των αποτελεσμάτων απόδοσης και των δύο τεχνικών. Τέλος στο έκτο κεφάλαιο αναλύονται τα εξαγόμενα συμπεράσματα από τα αποτελέσματα απόδοσης των δύο μεθοδολογιών ενώ επίσης εξετάζονται οι προοπτικές εξέλιξης της παρούσας διπλωματικής.

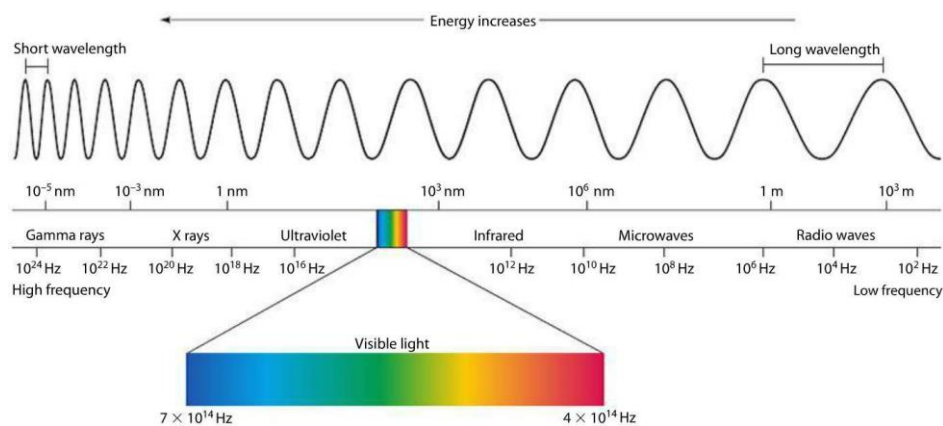
Κεφάλαιο 2^ο

Το κεφάλαιο αυτό παρέχει μια επισκόπηση των θεμελιωδών αρχών της υπέρυθρης απεικόνισης, συμπεριλαμβανομένων των αρχών της φυσικής που διέπουν την ακτινοβολία, της τεχνολογίας των θερμικών καμερών και απεικόνισης των αερίων, καθώς και μία σύνοψη των υφιστάμενων τεχνολογιών ανίχνευσης αερίων.

2 Υπέρυθρη θερμική απεικόνιση

2.1 Υπέρυθρη Ακτινοβολία

Με τον όρο ηλεκτρομαγνητικό φάσμα περιγράφεται το εύρος της περιοχής συχνοτήτων που καλύπτουν τα ηλεκτρομαγνητικά κύματα, τα οποία θεωρητικά μπορούν να έχουν από μηδενικές τιμές έως το άπειρο [4]. Σύμφωνα με ορισμένες χαρακτηριστικές ιδιότητες των ηλεκτρομαγνητικών κυμάτων το ηλεκτρομαγνητικό φάσμα κατηγοριοποιείται σε επιμέρους ζώνες, οι οποίες είναι οι ακτίνες γ, ακτίνες Χ, η υπεριώδης ακτινοβολία, η ορατή ακτινοβολία, η υπέρυθρη ακτινοβολία, τα μικροκύματα και τα ραδιοκύματα.



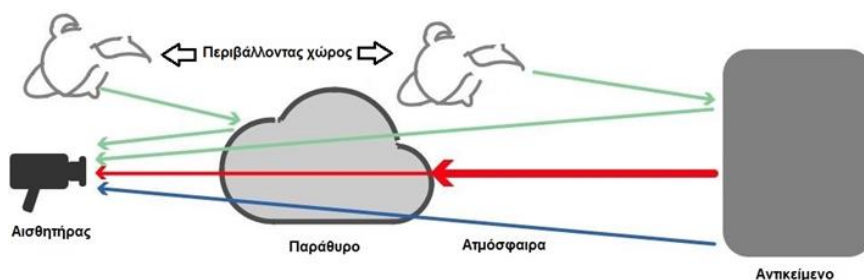
Εικόνα 2.1: Το ηλεκτρομαγνητικό φάσμα και οι επιμέρους κατηγορίες του [5].

Συνεπώς, όπως φαίνεται και παραπάνω, η ανθρώπινη αντίληψη για το ηλεκτρομαγνητικού φάσματος είναι μάλλον περιορισμένη. Το υπέρυθρο ηλεκτρομαγνητικό φάσμα χωρίζεται σε τρεις υποκατηγορίες: στο κοντινό (near IR, 0,70-1,50μm), στο ενδιάμεσο (middle IR, 1,50-5,60μm) και στο μακρινό (far IR, 5,60-1.000μm) [6]. Πρέπει να σημειωθεί εδώ πως οι ορισμοί και οι οριοθετήσεις των μηκών κύματος δεν είναι ίδιες σε όλη τη βιβλιογραφία καθώς δίνονται διαφορετικές εκδοχές στην κατηγοριοποίηση του υπέρυθρου φάσματος.



Εικόνα 2.2: Διαρροή προπάνιου που λήφθηκε με τη θερμική κάμερα Ventus της Sierra Olympic.

Οι θερμικές κάμερες, όπως αυτές της ανίχνευσης αερίων, λειτουργούν κυρίως στις περιοχές 3-5μm και 10-11μm του υπέρυθρου φάσματος λόγω των μέγιστων απορροφήσεων ακτινοβολίας των αερίων στις συγκεκριμένες φασματικές περιοχές. Το φασματικό εύρος των θερμικών καμερών επηρεάζεται επίσης από τα χαρακτηριστικά του ανιχνευτή της κάμερας και την αλληλεπίδραση της ακτινοβολίας με την ύλη. Η παραπάνω εικόνα, που λήφθηκε με θερμική κάμερα στο υπέρυθρο, απεικονίζει τη διαρροή προπάνιου σε ένα διωλιστήριο. Με γνώμονα το παράδειγμα αυτό, γίνεται αντιληπτό ότι οι εικόνες υπέρυθρης ακτινοβολίας παρέχουν πληροφορίες τόσο για το χωρικό όσο και για το ενεργειακό περιεχόμενο της σκηνής που συλλαμβάνουν. Η ηλεκτρομαγνητική ακτινοβολία διέρχεται τη γήινη ατμόσφαιρα και φτάνει στην επιφάνεια με την οποία αντιδρά με ανάκλαση-διάχυση, απορρόφηση ή με μετάδοση.



Εικόνα 2.3: Η ακτινοβολία του αντικειμένου (κόκκινη γραμμή), η ακτινοβολία της ατμόσφαιρας (μπλε γραμμή) και η ανάκλαση της ακτινοβολίας από τον περιβάλλοντα χώρο (πράσινη γραμμή) [4].

Στη περίπτωση της ανάκλασης, μία ακτίνα φωτός ανακατευθύνεται καθώς προσπίπτει σε μια αδιάφανη επιφάνεια, ενώ η φύση της αντανάκλασης εξαρτάται από τα μεγέθη των επιφανειακών ανωμαλιών (τραχύτητα ή ομαλότητα). Η περίπτωση της μετάδοσης της ακτινοβολίας συμβαίνει όταν η ηλιακή ακτινοβολία περνάει μέσα από μία ουσία χωρίς σημαντική εξασθένηση, ενώ ο όρος διαπερατότητα περιγράφει την ικανότητα ενός μέσου για τη μετάδοση της ενέργειας για ένα δεδομένο πάχος ή βάθος μιας ουσίας. Τέλος, η απορρόφηση εκφράζει το φαινόμενο κατά το οποίο η ακτινοβολία δεσμεύεται από το αντικείμενο και ακολούθως επανεκπέμπεται σε μεγαλύτερα μήκη κύματος. Η ακτινοβολία διάφορων στοιχείων έξω από το οπτικό πεδίο μίας θερμικής κάμερας μπορεί να επηρεάσει την εικόνα

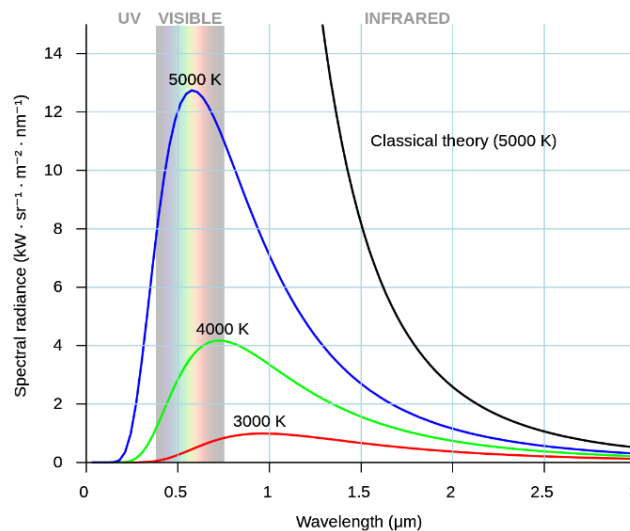
αντανακλώντας είτε στο εικονιζόμενο αντικείμενο, είτε σε οποιοδήποτε άλλο αντικείμενο κατά μήκος της οπτικής διαδρομής. Συνεπώς η φαινομενική θερμοκρασία του αντικειμένου, δεν καθορίζεται μόνο από την πραγματική θερμοκρασία του, αλλά από έναν αριθμό περιβαλλοντικών παραγόντων. Οι απώλειες της ενέργειας της ακτινοβολίας κατά την απορρόφηση, την κατοπτρική ανάκλαση, τη διάχυση, καθώς και την μετάδοση ακτινοβολίας μέσω του αντικειμένου χαρακτηρίζονται από τρεις συντελεστές, τον συντελεστή απορρόφησης $\alpha(\lambda)$, το συντελεστή ανακλάσεως $\rho(\lambda)$ και το συντελεστή διελεύσεως $\tau(\lambda)$ αντίστοιχα, το άθροισμα των οποίων ισούται πάντα με τη μονάδα:

$$\alpha(\lambda) + \rho(\lambda) + \tau(\lambda) = 1 \quad (2.1.1)$$

Δεδομένου ότι η υπέρυθη ακτινοβολία αφορά σε μία συγκεκριμένη περιοχή του ηλεκτρομαγνητικού φάσματος, κρίνεται απαραίτητος ο προσδιορισμός του μεγέθους της ηλεκτρομαγνητικής ακτινοβολίας για συγκεκριμένα μήκη κύματος. Ο υπολογισμός αυτού του ποσού της ηλεκτρομαγνητικής ακτινοβολίας που ακτινοβολείται από ένα σώμα για συγκεκριμένη θερμοκρασία και μήκος κύματος υπολογίζεται από το νόμο του Planck:

$$E(T) = c_1 * \lambda^{-5} * e^{\frac{c_2}{\lambda T}} * d\lambda \quad (2.1.2)$$

,όπου $c_1=3,74*10^{-16} \text{ W/m}^2$, $c_2=1,44*10^{-2} \text{ mK}$, λ =μήκος κύματος και T =απόλυτη θερμοκρασία (σε Kelvin).



Εικόνα 2.4: Κατανομή της ακτινοβολίας μαύρου σώματος για διαφορετικές θερμοκρασίες σύμφωνα με τον νόμο του Planck [7].

Η σχέση μεταξύ ενεργειακών ποσοτήτων και θερμοκρασίας διατυπώνεται από τον νόμο των Stephan-Boltzman, ο οποίος αποτελεί ολοκλήρωμα του νόμου του Planck για όλα τα μήκη κύματος της ακτινοβολίας. Ο νόμος των Stephan-Boltzman ορίζει πως η συνολικά εκπεμπόμενη ροή ενέργειας (E σε watts/m^2) του μέλανος σώματος για όλα τα μήκη κύματος της ακτινοβολίας είναι ανάλογη της τέταρτης δύναμης της απόλυτης θερμοκρασίας (T) του σώματος και μετριέται σε βαθμούς Kelvin:

$$E = \sigma T^4 \quad (2.1.3)$$

,όπου σ η σταθερά των Stephan-Boltzman που ισούται με:

$$\sigma = 5,669 * 10^{-8} \text{ Wcm}^{-2} \text{ K}^{-4} \quad (2.1.4)$$

Για ένα μη μέλαν σώμα στη σχέση (2.1.3) εισέρχεται ένα κλάσμα ε το οποίο ορίζει τη συνολική ροή που εκπέμπεται από το μέλαν σώμα ίδιας θερμοκρασίας και καλείται ικανότητα εκπομπής. Το μέγιστο μήκος κύματος στην εξίσωση του Planck δίνεται από το νόμο του Wien που ορίζει ότι το μήκος κύματος, κατά το οποίο παρατηρείται η μέγιστη εκπομπή φωτός από μια φωτεινή πηγή, είναι αντιστρόφως ανάλογο της θερμοκρασίας της πηγής:

$$\lambda_{\max} = \frac{2898 \text{ K}\mu\text{m}}{T} \quad (2.1.5)$$

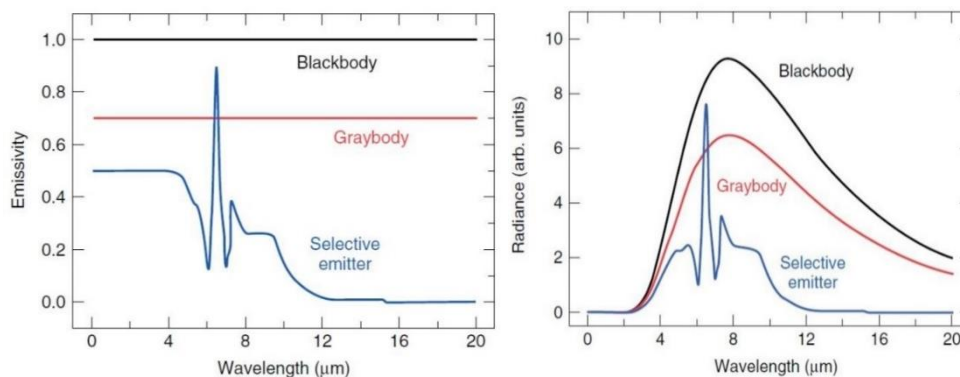
Η έννοια του μέλανος σώματος ουσιαστικά αποτελεί ένα εξιδανικευμένο μοντέλο της ύλης το οποίο απορροφά όλη την προσπίπτουσα σε αυτό ηλεκτρομαγνητική ακτινοβολία ανεξαρτήτως συχνότητας και ανεξαρτήτως γωνίας προσπτώσεως. Σύμφωνα με τον νόμο θερμικής ακτινοβολίας του Kirchhoff, σε θερμική ισορροπία, η ισχύς που ακτινοβολείται από ένα αντικείμενο $\varepsilon(\lambda)$ πρέπει να είναι ίση με την απορροφημένη ισχύ $\alpha(\lambda)$ (ενεργειακό ισοζύγιο). Συνεπώς ένα μέλαν σώμα ανακλά μηδενικό ποσοστό της προσπίπτουσας ακτινοβολίας ($\rho = 0$), εσωτερικά απορροφά όλη αυτή την προσπίπτουσα ακτινοβολία ($\alpha = 1$), χωρίς να αφήνει ακτινοβολία να το διαπεράσει ($\tau = 0$), για όλα τα μήκη κύματος και για όλες τις γωνίες προσπτώσεως. Επιπλέον η εξίσωση (2.1.1), μπορεί να αναδιατυπωθεί :

$$\varepsilon(\lambda) + \rho(\lambda) + \tau(\lambda) = 1 \quad (2.1.6)$$

Επομένως, ένα ιδανικό μέλαν σώμα έχει ικανότητα εκπομπής ε ίση με τη μονάδα. Για κάθε πραγματικό σώμα όμως ο συντελεστής εκπομπής είναι μικρότερος από την μονάδα, όπου ως συντελεστής εκπομπής ορίζεται ο λόγος της ακτινοβολούμενης ενέργειας από το σώμα σε σχέση με την ακτινοβολούμενη ενέργεια ενός μελανού σώματος που βρίσκεται στην ίδια θερμοκρασία:

$$\varepsilon(\lambda) = \frac{E_{\text{real}}(\Delta\lambda, T)}{E_{\text{bb}}(\Delta\lambda, T)} \quad (2.1.7)$$

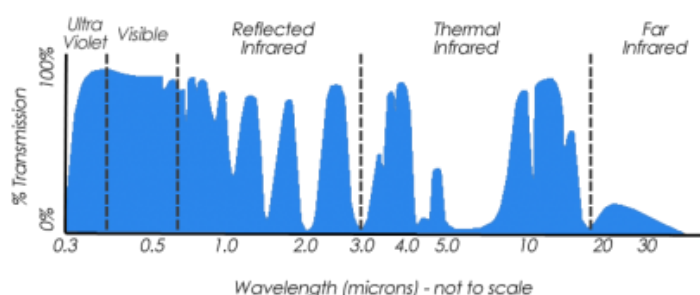
,όπου E_{real} και E_{bb} τα ποσά ακτινοβολίας ενός πραγματικού σώματος και ενός μελανού σώματος αντίστοιχα. Για τα περισσότερα αντικείμενα, γνωστά ως γκρι σώματα, η τιμή εκπομπής είναι σταθερή σε δεδομένη θερμοκρασία. Ωστόσο, εάν ένα αντικείμενο παρουσιάζει ζώνες απορρόφησης και εκπομπής στο υπέρυθρο ηλεκτρομαγνητικό φάσμα, όπως συμβαίνει με το νερό και με τα αέρια, η ικανότητα εκπομπής μεταβάλλεται συναρτήσει του μήκους κύματος. Ένα τέτοιο αντικείμενο χαρακτηρίζεται ως επιλεκτικός πομπός, αφού λειτουργεί ως πηγή η οποία ακτινοβολεί σε επιλεκτικά μήκη κύματος. Η παρακάτω εικόνα απεικονίζει τις τιμές εκπομπής για του μέλανος σώματος, του γκρι σώματος και ενός επιλεκτικού πομπού σε συναρτήση του μήκους κύματος.



Εικόνα 2.5: Εκπομπή (αριστερά) και η θερμική ακτινοβολία (δεξιά) ως συνάρτηση του μήκους κύματος ενός μελανού σώματος, ενός γκρι σώματος και ενός επιλεκτικού πομπού [8].

2.2 Θερμικές Κάμερες

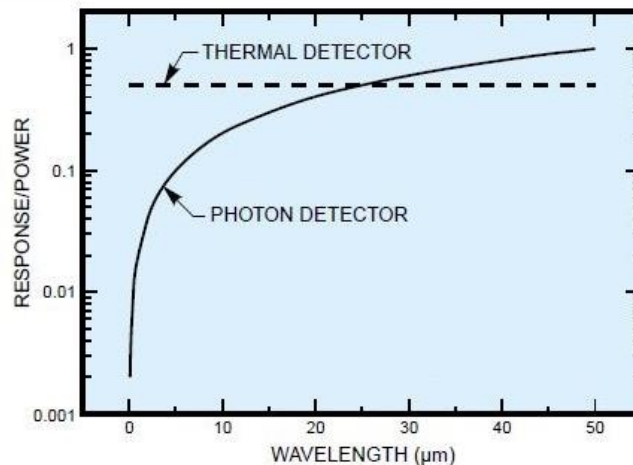
Οι συμβατικές κάμερες καταγράφουν την ηλεκτρομαγνητική ακτινοβολία μιας σκηνής στο ορατό φάσμα ($0,38\mu\text{m}$ με $0,78\mu\text{m}$) που ανακλάται, μεταδίδεται και διασκορπίζεται από τα αντικείμενα που υπάρχουν στην σκηνή. Ο αισθητήρας μιας τέτοιας κάμερας επιτρέπει την καταγραφή χρωμάτων, ενώ ορισμένοι μπορούν να καταγράψουν ηλεκτρομαγνητική ακτινοβολία από ένα μικρό κομμάτι του υπέρυθρου φάσματος στο εγγύς υπέρυθρο που βρίσκεται δίπλα στο ορατό φάσμα. Σε αντίθεση, οι θερμικές, ή αλλιώς υπέρυθρες, κάμερες καταγράφουν την ακτινοβολία που εκπέμπεται από τα ίδια τα αντικείμενα που απεικονίζονται, διαμορφώνοντας μια αναπαράσταση της ενεργειακής δομής της σκηνής. Όπως έχει αναφερθεί, η απεικόνιση μιας σκηνής στο υπέρυθρο φάσμα επηρεάζεται και από την ατμόσφαιρα, η οποία προκαλεί εξασθένηση επιλεκτικά της ακτινοβολίας ανάλογα με το μήκος κύματός της. Οι περιοχές του υπέρυθρου φάσματος που επιτρέπουν σε συγκεκριμένα μήκη κύματος της ηλεκτρομαγνητικής ακτινοβολίας ανεμπόδιστα να διαπεράσουν την ατμόσφαιρα, χωρίς ενδιάμεση απορρόφηση, ονομάζονται ατμοσφαιρικά παράθυρα.



Εικόνα 2.6: Τα ατμοσφαιρικά παράθυρα στο ηλεκτρομαγνητικό φάσμα [9].

Τα ατμοσφαιρικά παράθυρα δεν μπορούν να οριοθετηθούν κατά μήκος του ηλεκτρομαγνητικού φάσματος, καθώς η φασματική σύνθεση της ακτινοβολίας του παραθύρου ποικίλλει σημαντικά ανάλογα με τις τοπικές περιβαλλοντικές συνθήκες [10]. Συνεπώς, το ατμοσφαιρικό παράθυρο ορίζεται ως μια δυναμική ιδιότητα της ατμόσφαιρας, ενώ αντιθέτως, το φασματικό παράθυρο είναι ένα στατικό χαρακτηριστικό απορρόφησης ηλεκτρομαγνητικής ακτινοβολίας που παρουσιάζουν ορισμένα αέρια, κυρίως τα αέρια του θερμοκηπίου, συμπεριλαμβανομένων των υδρατμών. Με γνώμονα τα παραπάνω, οι περισσότερες θερμικές κάμερες, προκειμένου να αποφύγουν την απώλεια σήματος μέσω της ατμοσφαιρικής απορρόφησης, χρησιμοποιούν ανιχνευτές με απόκριση συνήθως στο θερμικό υπέρυθρο φάσμα ($8\mu\text{m}$ - $12\mu\text{m}$) ή στο κοντινό υπέρυθρο φάσμα ($2\mu\text{m}$ - $5\mu\text{m}$) [11].

Επομένως, η χρήση μιας θερμικής κάμερας που έχει απόκριση σε συγκεκριμένη περιοχή του υπέρυθρου φάσματος, σχετίζεται άμεσα με τις απαιτήσεις της εφαρμογής, όπως για παράδειγμα, σε πολλά αέρια που εμφανίζουν υψηλά ποσοστά απορρόφησης στην περιοχή $3\mu\text{m}$ - $4\mu\text{m}$ του ηλεκτρομαγνητικού φάσματος. Εκτός από την χρήση ενός ανιχνευτή με συγκεκριμένη απόκριση, οι νέες κάμερες αντισταθμίζουν την επιρροή της ατμόσφαιρας χρησιμοποιώντας υπολογιστικά μοντέλα διόρθωσης.



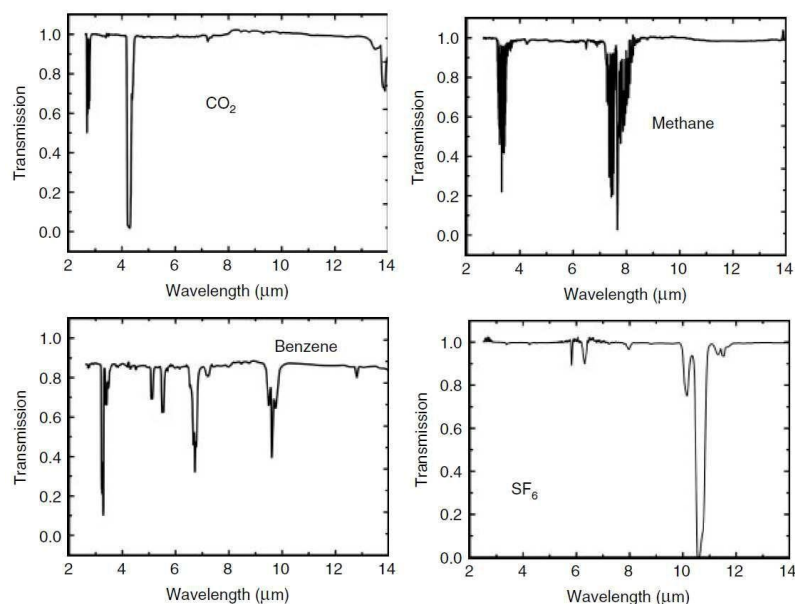
Εικόνα 2.7: Σχετικές φασματικές αποκρίσεις ανιχνευτών φωτονίων και θερμικών ανιχνευτών [12].

Μία θερμική κάμερα, ή αλλιώς μια μονάδα ανίχνευσης, μετατρέπει τα ποσά θερμικής ενέργειας που ανιχνεύει ο αισθητήρας σε ηλεκτρικά σήματα. Οι θερμικές κάμερες που διαχωρίζονται με βάση τον αισθητήρα ανίχνευσης που χρησιμοποιούν σε δύο κατηγορίες, στους ανιχνευτές φωτονίων και στους θερμικούς ανιχνευτές. Ο αισθητήρας ανίχνευσης κατηγοριοποιείται επίσης ανάλογα και με τον τρόπο παραγωγής του ηλεκτρικού σήματος. Οι ανιχνευτές φωτονίων μετατρέπουν την απορροφημένη ηλεκτρομαγνητική ακτινοβολία απευθείας σε μια αλλαγή της κατανομής της ηλεκτρικής ενέργειας σε έναν ημιαγωγό με την αλλαγή της συγκέντρωσης του ελεύθερου φορέα φόρτισης. Αντίθετα, οι θερμικοί ανιχνευτές μετατρέπουν την απορροφημένη ηλεκτρομαγνητική ακτινοβολία σε θερμική ενέργεια προκαλώντας αύξηση της θερμοκρασίας του ανιχνευτή. Στη συνέχεια, η ηλεκτρική έξοδος του θερμικού αισθητήρα παράγεται από μια αντίστοιχη αλλαγή σε κάποια φυσική ιδιότητα ενός υλικού, πχ τη θερμοκρασιακή διαφορά της αντίστασης ενός βολτόμετρου. Συνήθως, οι ανιχνευτές φωτονίων λειτουργούν στο μέση ζώνη του υπέρυθρου φάσματος, όπου η θερμική αντίθεση είναι υψηλή, καθιστώντας τον ανιχνευτή ευαίσθητο στις μικρές διαφορές θερμοκρασίας της σκηνής, ενώ ταυτόχρονα παρουσιάζουν υψηλούς χρόνους απόκρισης. Το κύριο μειονέκτημα ενός τέτοιου ανιχνευτή είναι η ανάγκη ψύξης, καθώς πρέπει να ψύχεται σε κρυογονικές θερμοκρασίες μικρότερες των 77 Kelvin προκειμένου να επιτευχθούν χαμηλά θερμικά επίπεδα θορύβου. Από την άλλη πλευρά, οι θερμικοί ανιχνευτές απαιτούν μόνο σταθεροποίηση της θερμοκρασίας αντί για την πλήρη ψύξη αλλά χαρακτηρίζονται από αργούς χρόνους απόκρισης και από χαμηλή ευαισθησία. Ένα σύστημα ανιχνευτή φωτονίων έχει υψηλότερο κόστος σε σχέση με ένα θερμικού ανιχνευτή τόσο λόγω του υψηλού κόστους κτήσης όσο και του κόστους συντήρησης. Το ποσό της θερμικής ενέργειας ενός αντικείμενου υπολογίζεται από μία θερμική κάμερα σύμφωνα με την καταγραφόμενη ενέργεια στη περιοχή του φάσματος στην οποία λειτουργεί η κάμερα και σε σύγκριση με την ενέργεια που ακτινοβολεί το μέλαν σώμα για μία συγκεκριμένη θερμοκρασία, σύμφωνα με το νόμο του Planck (2.1.2). Στη περίπτωση που είναι γνωστό το ποσό εκπομπής ακτινοβολίας από το αντικείμενο, μπορεί να υπολογισθεί η πραγματική θερμοκρασία, διαφορετικά η κάμερα υποθέτει μια προκαθορισμένη τιμή εκπομπής. Οι υποτιθέμενες ή υπολογισμένες θερμοκρασίες όλων των στοιχείων μιας σκηνής, εμφανίζονται στην οθόνη της κάμερας είτε με κλίμακα του γκρι, που οι σκούρες και φωτεινές περιοχές αντιστοιχούν σε χαμηλές και υψηλές θερμοκρασίες αντίστοιχα, ή χρησιμοποιώντας ένα πίνακα αντιστοίχισης ψευδοχρώματος. Όπως φαίνεται και στην Εικόνα 2.2, το αέριο είναι ορατό ως μαύρο σύννεφο καπνού, γεγονός που καταδεικνύει ότι η φαινομενική θερμοκρασία του αερίου είναι χαμηλότερη από αυτή του υποβάθρου.

2.3 Θερμογραφία Υπέρυθρου στα Αέρια

Όπως αναφέρθηκε, σε μια θερμική εικόνα το αέριο εμφανίζει χαμηλότερη φαινομενική θερμοκρασία σε σχέση με τη θερμοκρασία υποβάθρου καθώς εμφανίζεται πιο σκούρο. Όσο αφορά την απεικόνιση ενός στερεού αντικείμενου με μία θερμική κάμερα, η θερμοκρασιακή διαφορά μεταξύ αντικείμενου και υποβάθρου δεν προκαλείται κυρίως από την εκπεμπόμενη ακτινοβολία του, αλλά από την απορρόφηση της ακτινοβολίας από το αντικείμενο. Η ικανότητα εκπομπής ή αντίστοιχα, η ικανότητα απορρόφησης των αερίων ποικίλει στα διαφορετικά τμήματα του ηλεκτρομαγνητικού φάσματος. Σε ορισμένα μήκη κύματος, τα αέρια εμφανίζουν ισχυρές γραμμές απορρόφησης σε αυτά τα μήκη κύματος, ενώ η θέση και το σχήμα αυτών των γραμμών καθορίζεται από έναν αριθμό παραγόντων σε μοριακό επίπεδο.

Για παράδειγμα, το πολυατομικό μόριο του μεθανίου, σχηματίζεται από τέσσερα άτομα υδρογόνου τα οποία είναι ικανά να απορροφούν, να εκπέμπουν και να διασκορπίζουν φωτόνια. Όταν η ενέργεια των φωτονίων αντιστοιχεί με την ποσοτικοποιημένη ενεργειακή διαφορά μεταξύ των δύο ενεργειακών καταστάσεων τότε το άτομο διεγείρεται από το φωτόνιο και μεταπηδά ενεργειακή στάθμη. Η απορρόφηση της ακτινοβολίας στην περιοχή του υπέρυθρου φάσματος προκαλεί διεγέρσεις μεταξύ διάφορων ενεργειακών σταθμών δόνησης και περιστροφής του μορίου, πάνω και γύρω από τους άξονες σύνδεσης, ενώ το μόριο παραμένει στη θεμελιώδη ηλεκτρονιακή του κατάσταση. Οι ενεργειακές δονήσεις του μορίου κατατάσσονται σε δύο κατηγορίες, στις δονήσεις τάσεως και στις δονήσεις κάμψης, όπου στις πρώτες η δόνηση πραγματοποιείται κατά μήκος του άξονα του δεσμού αλλάζοντας την απόσταση των συνδεόμενων δονούμενων ατόμων, ενώ οι δεύτερες αφορούν στη δόνηση που επιδρά στη γεωμετρία του μορίου όπου μεταβάλλονται οι γωνίες των μοριακών δεσμών.



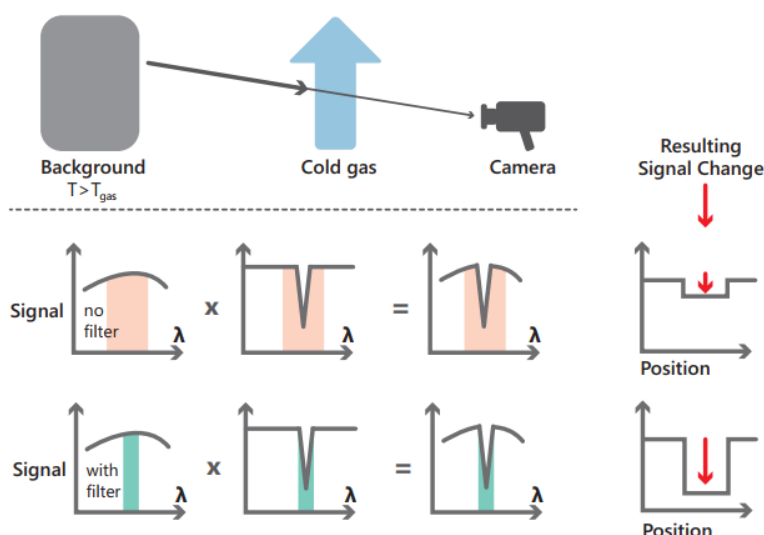
Εικόνα 2.8: Το ποσοστό μετάδοσης ακτινοβολίας των αερίων του διοξειδίου του άνθρακα, του μεθανίου, του βενζολίου και του εξαφθοριούχου θείου αντίστοιχα [4].

Παρόλα αυτά, η υπέρυθρη ακτινοβολία δεν διαθέτει αρκετή ενέργεια για να προκαλέσει τα είδη των ηλεκτρονιακών μεταπτώσεων που συναντώνται στην υψηλότερης ενέργειας ορατή και υπεριώδη ακτινοβολία. Αντίθετα, οι ενέργειες που απαιτούνται για τις μεταβατικές δονήσεις βρίσκονται εντός του υπέρυθρου φάσματος. Σύμφωνα με την αρχή της κβαντομηχανικής, τον *κανόνα επιλογής*, εάν ένα μόριο που βρίσκεται σε μια κατώτερη ενεργειακή στάθμη απορροφήσει μια ποσότητα ακτινοβολίας, θα διεγερθεί και θα μεταπηδήσει

σε υψηλότερο ενεργειακό επίπεδο, ακολουθώντας τις επιτρεπτές μεταβάσεις που καθορίζει η κβαντομηχανική. Στην περίπτωση που οι δονήσεις ενός μορίου περιλαμβάνουν αλλαγές στην διπολική ροπή του, τότε το μόριο αλληλοεπιδρά στα υπέρυθρα τμήματα και ανάλογα με τη μετάβαση, προκαλεί απορρόφηση ή εκπομπή ενέργειας, η οποία μπορεί να καταγραφεί. Μόρια που εμφανίζουν ανισοκατανομή των ηλεκτρονίων, όπως είναι τα ετεροπυρηνικά διατομικά μόρια, εμφανίζουν μια εναλλασσόμενη διπολική ροπή, που καθορίζει την χημική πολικότητα ενός χημικού δεσμού μέσα σε ένα μόριο και η οποία εξαρτάται από την ένταση του φορτίου και την απόσταση μεταξύ των κέντρων φόρτισης.

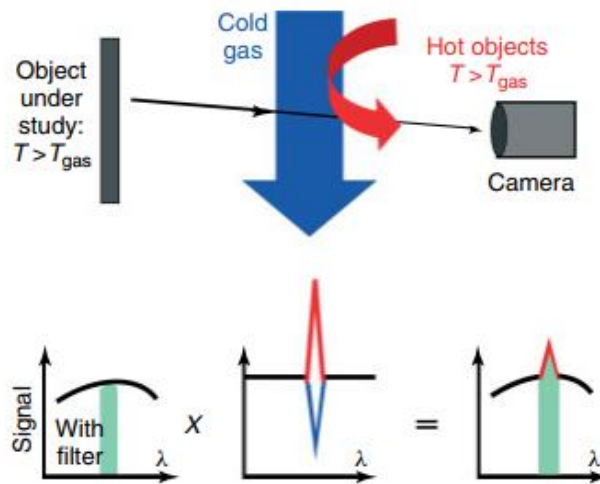
Κατά τη διάρκεια ενός μοριακού κραδασμού, η διπολική ροπή μπορεί να μεταβληθεί περιοδικά σε μία στιγμή. Εάν η συχνότητα της ταλάντωσης αυτής στη στιγμή της σύντομης περιοδικής αλλαγής ταιριάζει με τη συχνότητα ενός προσπίπτοντος φωτονίου, τότε το μόριο του αερίου μπορεί να απορροφήσει την ενέργεια του φωτονίου. Τα μονοατομικά αέρια όπως είναι το άζωτο N_2 και ή το οξυγόνο O_2 δεν διαθέτουν διπολική ροπή λόγω της ομοιόμορφης δομής και συνεπώς μπορούν να τα διεγείρουν τα φωτόνια.

Η χαρτογράφηση της απορρόφησης που προκαλείται από τους μοριακούς κραδασμούς, όπως φαίνεται και στην παραπάνω εικόνα, παρουσιάζει κορυφές των οποίων η θέση και η πολυπλοκότητα καθορίζονται από τον τύπο, τον αριθμό και το μέγεθος του μοριακού κραδασμού καθώς και από τις μεταπηδήσεις στα διαφορετικά ενεργειακά επίπεδα σύμφωνα με τον κανόνα επιλογής.



Εικόνα 2.9: Ανίχνευση ψυχρών αερίων μπροστά από ένα θερμό υπόβαθρο με ή χωρίς φίλτρο λόγω της απορρόφησης. Οι περιοχές με την κόκκινη και πράσινη σκιά αντικατοπτρίζουν την ανιχνευθείσα φασματική περιοχή του συστήματος της κάμερας [4].

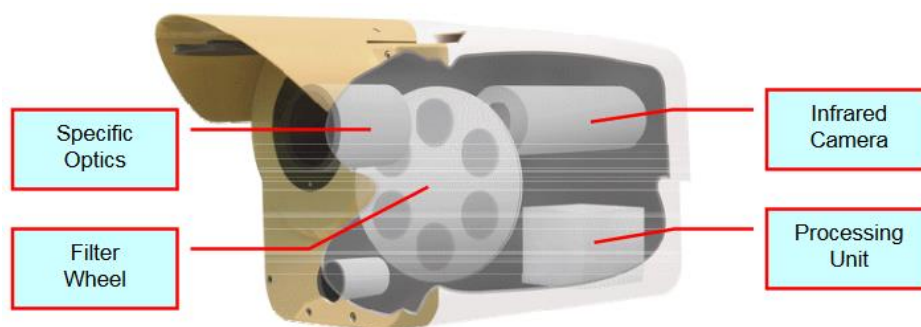
Υπάρχουν τρεις διαφορετικές περιπτώσεις διάδρασης μεταξύ των φωτονίων και των μορίων του αερίου. Η πρώτη αφορά στην ικανότητα των μορίων του αερίου όταν βρίσκονται στη θεμελιώδη ενεργειακή στάθμη να εμφανίζουν υψηλή απορροφητικότητα προς την υπέρυθρη ακτινοβολία. Στην περίπτωση αυτή, το ποσό της υπέρυθρης ακτινοβολίας ενός αντικειμένου που βρίσκεται πίσω από το αέριο που ανιχνεύεται είναι μικρότερη σε σχέση με την ακτινοβολία που θα ανιχνευόταν εάν απουσίαζε το αέριο.



Εικόνα 2.10: Αύξηση σήματος λόγω της διάχυσης της υπέρυθρης ακτινοβολίας από ένα θερμό αντικείμενο στα μόρια του αερίου [4].

Η δεύτερη αφορά στο φαινόμενο που η υπέρυθρη ακτινοβολία μπορεί να εκπέμπεται ισότροπα εάν τα μόρια του αερίου έχουν μεταπηδήσει ενεργειακή στάθμη, βρίσκονται δηλαδή σε διέγερση. Σε αυτή την περίπτωση το ποσό της ανιχνευθείσας ενέργειας ενός αντικειμένου που βρίσκεται ξανά πίσω από το αέριο είναι μεγαλύτερη. Τέλος η ακτινοβολία μπορεί να διαχυθεί από τα μόρια του αερίου. Αυτό συμβαίνει όταν τα μόρια αρχικά βρίσκονται στη θεμελιώδη ενεργειακή στάθμη και στη συνέχεια μεταπηδούν ενεργειακή βαθμίδα λόγω της υπέρυθρης ακτινοβολίας θερμότερων αντικειμένων που βρίσκονται κοντά.

Αμέσως μετά την διέγερση μεταδίδουν πίσω αυτή την υπέρυθρη ακτινοβολία που συνετέλεσε στην αλλαγή της ενεργειακής στάθμης τους σχεδόν ισότροπα. Ξανά η ανιχνευθείσα ακτινοβολία θα είναι υψηλότερη σε σχέση με την απουσία του αερίου. Σε μία πειραματική εφαρμογή όπου επιχειρείται η απεικόνιση της υπέρυθρης ακτινοβολίας μιας σκηνής και χρησιμοποιείται ένα φασματικό φίλτρο στενής ζώνης (bandpass filter) μπροστά από τον ανιχνευτή της κάμερας, το σήμα του αντικειμένου είναι μικρότερο από εκείνο που επιτυγχάνεται χωρίς την εφαρμογή φίλτρου. Από την άλλη, εάν το φίλτρο αφορά συγκεκριμένα στο μήκος κύματος όπου ένα συγκεκριμένο αέριο εμφανίζει το υψηλότερο ποσοστό απορρόφησης του, η σχετική αυτή διαφορά του σήματος θα δώσει υψηλότερες τιμές αντίθεσης και συνεπώς αυξάνεται η ευαισθησία της ανίχνευσης.



Εικόνα 2.11: Σύστημα θερμικής κάμερας ανίχνευσης αερίων [13].

Γενικά, η οπτική απεικόνιση ενός αερίου με μία απλή κάμερα υπέρυθρων είναι εφικτή μόνο στην περίπτωση που εκείνη διαθέτει ανιχνευτή του οποίου το εύρος ζώνης συμπεριλαμβάνει τις ισχυρές γραμμές απορρόφησης του αερίου. Βέβαια, καθώς το αέριο ανιχνεύεται σε ένα

συγκεκριμένο εύρος μηκών κύματος (π.χ. 3.3-3.5 μm) χρησιμοποιώντας ολόκληρο το φάσμα του ανιχνευτή (π.χ. 2-5 μm) η αντίθεση της θερμικής απεικόνισης μεταξύ αερίου και υποβάθρου θα είναι πολύ μικρή. Ως εκ τούτου, οι περισσότερες θερμικές κάμερες που κατασκευάζονται για την ανίχνευση αερίων, διαθέτουν το προαναφερθέν φασματικό φίλτρο στενής ζώνης μπροστά από τον αισθητήρα της κάμερας, προκειμένου να επιτευχθεί η μέγιστη μείωση της ισχύος του σήματος υποβάθρου και συνεπακόλουθα της υψηλής αντίθεσης σήματος αερίου-στόχου και υποβάθρου.

Κεφάλαιο 3^ο

Στο κεφάλαιο αυτό αναλύονται οι θεμελιώδεις αρχές ανίχνευσης αντικειμένων σε μία εικόνα. Αρχικά γίνεται μια συνοπτική παρουσίαση των τεχνικών προεπεξεργασίας και αφαίρεσης υποβάθρου καθώς και μια γενική ανάλυση των διαδικασιών εξαγωγής χαρακτηριστικών. Στη συνέχεια αναλύονται λεπτομερώς οι στατικοί και δυναμικοί περιγραφείς χαρακτηριστικών, οι αλγόριθμοι ταξινόμησης που χρησιμοποιήθηκαν καθώς και οι αλγόριθμοι εξαγωγής, αντιστοίχισης χαρακτηριστικών και ψηφιακής επεξεργασίας εικόνας που υλοποιήθηκαν για την ανάπτυξη της τεχνικής εξαγωγής υποβάθρου. Στο τέλος του κεφαλαίου παρατίθεται αναλυτικά η βιβλιογραφία των υπαρχόντων συστημάτων θερμικής απεικόνισης αερίων και συστημάτων ανίχνευσης καπνού σε βίντεο.

3 Ανίχνευση στα στιγμιότυπα των βίντεο

3.1 Προεπεξεργασία εικόνας και αφαίρεση υποβάθρου

Η προεπεξεργασία της εικόνας πραγματοποιείται συνήθως για να βελτιωθεί η ποιότητα των ακατέργαστων δεδομένων προκειμένου να γίνει πιο επιτυχής η εξαγωγή των χαρακτηριστικών. Στην ανίχνευση αντικειμένων στα στιγμιότυπα ενός βίντεο, η προεπεξεργασία αφορά διαδικασίες επιλογής υποδειγματοληψίας στιγμιότυπων, μετασχηματισμούς χρωμάτων στην κλίμακα του γκρι, τεχνικές ενίσχυσης, αντίθεσης, διαίρεσης των στιγμιότυπων σε δισδιάστατα ή τρισδιάστατα τμήματα, ή εξαγωγή της κίνησης των αντικειμένων. Υπάρχει ένας μεγάλος αριθμός αλγορίθμων ανίχνευσης κίνησης, όπου ο καθένας έχει τα πλεονεκτήματα και τα μειονεκτήματα του και εφαρμόζονται ανάλογα με το πεδίο της έρευνας. Η αφαίρεση του υποβάθρου (Background Subtraction) είναι μια ευρέως χρησιμοποιούμενη τεχνική για την κατάτμηση της εικόνας ανάλογα με τις κινούμενες περιοχές, όπου σύμφωνα με αυτή οι στατικές σκηνές ανήκουν στο υπόβαθρο. Οι τεχνικές αφαίρεσης υποβάθρου αποδίδουν σε ικανοποιητικά για την ανίχνευση της κίνησης ακόμα και εάν ένα κινούμενο αντικείμενο παραμείνει σταθερό για κάποιο χρονικό διάστημα. Παρόλα αυτά παρουσιάζουν προβλήματα στην περίπτωση που παρατηρηθούν μεγάλες διαφοροποιήσεις στον φωτισμό ή εάν κάποιο μέρος που θεωρήθηκε πως ανήκει στο παρασκήνιο αρχίσει να κινείται.

3.1.1 Αφαίρεση υποβάθρου μέσω τεχνικής διαδοχικών στιγμιότυπων

Η τεχνική των διαδοχικών στιγμιότυπων αποτελεί μια αρκετά διαδεδομένη και απλή τεχνική διαχωρισμού του υποβάθρου από τα κινούμενα αντικείμενα σε ασπρόμαυρο βίντεο με στατική κάμερα που βασίζεται στην διαφορά δύο διαδοχικών στιγμιότυπων. Η μεθοδολογία είναι αρκετά απλή, καθώς εάν η τιμή της διαφοράς των τιμών του εικονοστοιχείου είναι μεγαλύτερη

της τιμής του κατωφλίου που ορίζει ο χρήστης, τότε παίρνει την τιμή 1 ενώ αντίθετα, εάν είναι μικρότερη τότε παίρνει την τιμή 0. Το αποτέλεσμα αυτής της τεχνικής είναι δυαδικές εικόνες στιγμιότυπων που τα άσπρα εικονοστοιχεία υποδηλώνουν την κίνηση αντικειμένου. Τα προβλήματα αυτή της τεχνικής είναι αρκετά, ιδιαίτερα με την επιλογή της κατάλληλης τιμής κατωφλίου καθώς μια μικρή τιμή θα επιτρέπει την ανίχνευση πολύ μικρών κινήσεων που μπορεί να είναι εσφαλμένες εκτιμήσεις που οφείλονται σε θόρυβο λόγω της κάμερας, σε αλλαγές φωτισμού, σε αέρα κ.ά. Από την άλλη πλευρά, η επιλογή μιας μεγάλης τιμής κατωφλίου, δεν επιτρέπει την ανίχνευση αλλαγών εάν για παράδειγμα ένα αντικείμενο κινείται με μικρή ταχύτητα. Η απλότητα της μεθόδου αυτής εγκυμονεί δύο μεγάλα μειονεκτήματα, αφενός τα αντικείμενα με ομοιόμορφη κατανομή της τιμής της πυκνότητας τους τα εσωτερικά τους εικονοστοιχεία ερμηνεύονται ως μέρη του υποβάθρου και αφετέρου τα αντικείμενα αν παραμείνουν ακίνητα για χρονικό διάστημα ενός στιγμιότυπου ερμηνεύονται ξανά ως κομμάτι του υποβάθρου. Συνεπώς, με τη μέθοδο αυτή, από το κινούμενο αντικείμενο ανιχνεύεται κατά βάση το περίγραμμά του και όχι ολόκληρο το αντικείμενο. Κύριο χαρακτηριστικό αυτής της τεχνικής είναι ότι τα κινούμενα αντικείμενα αφήνουν πίσω τους ίχνη τα οποία ποικίλουν ανάλογα με την ταχύτητά τους.

3.1.2 Αφαίρεση υποβάθρου μέσω στατιστικών μεθόδων

Οι στατιστικές μέθοδοι είναι οι βασικότερες τεχνικές αφαίρεσης υποβάθρου. Οι στατιστικές μέθοδοι αναλύουν τα στοιχεία μεμονωμένων εικονοστοιχείων ή μιας ομάδας εικονοστοιχείων και χρησιμοποιούν τη στατιστική πληροφορία για να ταξινομήσουν περιοχές της εικόνας ως περιοχές που ανήκουν στο υπόβαθρο ή στο προσκήνιο. Η πιο δημοφιλής στατιστική μέθοδος είναι η γκαουσιανή που μοντελοποιεί κάθε εικονοστοιχείο και στη συνέχεια χρησιμοποιεί μια δυναμική διαδικασία προσέγγισης για την ενημέρωση του μοντέλου. Διαφορετικές στατιστικές μεθοδολογίες είναι η χρήση μεγίστων και ελαχίστων τιμών έντασης ή η χρήση της μέγιστης απόκλισης μεταξύ των μεγίστων και ελαχίστων τιμών για τον υπολογισμό και την μοντελοποίηση του υποβάθρου. Πιο συγκεκριμένα στη μοντελοποίηση υποβάθρου από μίξεις Gaussian, θεωρεί ως x μία τυχαία μεταβλητή που αντιπροσωπεύει την τιμή ενός δοθέντος εικονοστοιχείου στο χρόνο και χρησιμοποιεί μια παραμετρική γκαουσιανή κατανομή όπου η πυκνότητα αποτελείται από ένα άθροισμα γκαουσιανών κατανομών. Αν θεωρηθεί ότι η συνάρτηση πυκνότητας ενός μείγματος γκαουσιανών συμβολίζεται με $p(x)$ τότε ο τύπος της συνάρτησης ορίζεται ως:

$$p(x) = \sum_{k=1}^K w_k N(x; \mu_k, \sigma_k) \quad (3.1.1)$$

,όπου w_k είναι τα βάρη και $N(x; \mu_k, \sigma_k)$ είναι η κανονική πυκνότητα του μέσου μ_k .

3.1.3 Αφαίρεση υποβάθρου μέσω επιλεκτικά μεταβαλλόμενου μοντέλου

Η τεχνική του επιλεκτικά μεταβαλλόμενου μοντέλου είναι λίγο πιο σύνθετη από τις στατικές μεθόδους. Ο προσδιορισμός των κινούμενων περιοχών επιτυγχάνεται με την αφαίρεση της κάθε νέας εικόνας από την εικόνα του υποβάθρου που έχει ορισθεί. Η απόλυτη διαφορά του αποτελέσματος εάν ξεπερνά κάποιο προκαθορισμένο κατώφλι, τότε το αντίστοιχο εικονοστοιχείο θεωρείται πως αποτελεί εικονοστοιχείο του προσκήνιου. Προκειμένου το αρχικό μοντέλο να προσαρμόζεται στις σταδιακές αλλαγές, οι νέες πληροφορίες ενσωματώνονται με τη χρήση ενός απλού IIR φίλτρου το οποίο έχει την παρακάτω μορφή:

$$B_{n+1} = aI_n + (1 - a)B_n \quad (3.1.2)$$

, όπου a ο ρυθμός εκμάθησης (learning rate), που συνήθως έχει τιμή κοντά στο 0.05. Για να αποφευχθεί η εισαγωγή θορύβου στο μοντέλο από εικονοστοιχεία που δεν ανήκουν στο υπόβαθρο συνήθως εισάγεται και μια διαδικασία ανατροφοδότησης, όπου όταν ένα εικονοστοιχείο εκτιμάται πως ανήκει στο παρασκήνιο, τότε αναβαθμίζεται το μοντέλο στην επόμενη χρονική στιγμή σύμφωνα με τον παραπάνω τύπο, διαφορετικά αν ανήκει στο προσκήνιο, το υπόβαθρο μένει αμετάβλητο. Αυτή η επιπλέον διαδικασία αποτελεί για το μεγάλο πλεονέκτημα αυτής της μεθόδου καθώς με την ανατροφοδότηση του μοντέλου του υποβάθρου, τα εικονοστοιχεία που έχουν θεωρηθεί ότι ανήκουν στο προσκήνιο δεν επηρεάζουν τυχόν αλλαγές στο παρασκήνιο. Η τεχνική αυτή είναι ευαίσθητη στις αλλαγές φωτισμού, παρόλα αυτά ρυθμίζοντας κατάλληλα τον ρυθμό εκμάθησης a επιτυγχάνονται αξιόπιστα αποτελέσματα ακόμα και σε σκηνές με πολύ ευαίσθητες μεταβολές του φωτισμού.

3.2 Εξαγωγή Χαρακτηριστικών

Η διαδικασία εξαγωγής χαρακτηριστικών αναφέρεται στον εντοπισμό χαρακτηριστικών της εικόνας με σκοπό τον επιτυχή προσδιορισμό ενός αντικειμένου ενδιαφέροντος προκειμένου να μπορεί να διαχωρισθεί από τα υπόλοιπα αντικείμενα υποβάθρου. Ένα αντικείμενο μπορεί να είναι ένα πρόσωπο, ένα άτομο ή όπως αναλύεται στην παρούσα διπλωματική η ροή ενός αερίου. Η διαδικασία δημιουργίας περιγραφών χαρακτηριστικών (feature descriptors) είναι αρκετά σημαντική για την ανίχνευση αντικειμένων καθώς οι περιγραφείς είναι αυτοί που λειτουργούν ως βάση για τον αλγόριθμο ανίχνευσης, δεδομένου ότι παρέχουν στον ταξινομητή τις διακριτές πληροφορίες διαχωρισμού του αντικειμένου ενδιαφέροντος και των υπόλοιπων αντικειμένων. Πιο απλά, οι περιγραφείς εξάγουν πληροφορίες από τα εικονοστοιχεία μιας εικόνας ή των στιγμιότυπων ενός βίντεο και τις μετατρέπουν σε μικρότερης διάστασης περιγραφική πληροφορία που χρησιμοποιείται στην εκπαίδευση ενός ταξινομητή. Τα χαρακτηριστικά αυτά συνήθως αναφέρονται σε γραμμές, ακμές, κορυφογραμμές, γωνίες, σημεία ή και υποπεριοχές της εικόνας. Ένα παράδειγμα δημιουργίας περιγραφών χαρακτηριστικών για την ανίχνευση αντικειμένων είναι ο διαδεδομένος αλγόριθμος Haar-wavelet που χρησιμοποιείται στις περισσότερες ψηφιακές φωτογραφικές μηχανές για την ανίχνευση προσώπων, όπου αναγνωρίζει τα πρόσωπα από τις ξεχωριστές διαφορές των τιμών έντασης των εικονοστοιχείων γύρω από ορισμένες προκαθορισμένες περιοχές όπως είναι τα μάτια και τα φρύδια ή το στόμα. Ένας ακόμα διαδεδομένος περιγραφέας χαρακτηριστικών για την αναγνώριση ανθρώπων είναι ο περιγραφέας ιστογράμματος κλίσεων διαβαθμίσεων (Histogram of Orientated Features), ο οποίος χρησιμοποιείται συνήθως στην ανίχνευση πεζών στα συστήματα αυτόνομης οδήγησης των αυτοκινήτων καθώς αναγνωρίζει τους πεζούς στην όρθια στάση από την χαρακτηριστική σταθμισμένη κατανομή προσανατολισμού των άκρων τους (κεφάλι, κορμός, χέρια, πόδια). Αντίστοιχα, ο περιγραφέας τοπικών δυαδικών προτύπων (Local Binary Pattern), είναι ένα ισχυρό εργαλείο για την περιγραφή υφών σε μία εικόνα. Εκτός από τα προαναφερθέντα παραδείγματα που αφορούν σε εφαρμογές με στατικούς περιγραφείς, υπάρχει επίσης μια μεγάλη ποικιλία εφαρμογών με χρήση πιο περίπλοκων χωροχρονικών και δυναμικών περιγραφών, όπως η αναγνώριση μοτίβων ανθρώπινης συμπεριφοράς. Σε ορισμένες περιπτώσεις, είναι επιθυμητή η μείωση των διαστάσεων των εξαγόμενων περιγραφών χαρακτηριστικών προκειμένου να βελτιωθεί η απόδοση του ταξινομητή και να επιτευχθεί υψηλή απόδοση γενίκευσης. Υπάρχουν δύο μεθοδολογίες για την μείωση των διαστάσεων, αυτής της επιλογής χαρακτηριστικών και αυτής της εξαγωγής δεύτερης τάξης χαρακτηριστικών. Η πρώτη μέθοδος μειώνει το μέγεθος των εξαγόμενων χαρακτηριστικών με την επιλογή ενός υποσυνόλου χαρακτηριστικών, αποσκοπώντας στην αύξηση της γενίκευσης, όπου απομακρύνονται από το σύνολο εκπαίδευσης τα χαρακτηριστικά εκείνα που παρουσιάζουν χαμηλή διακύμανση μιας και είναι σχεδόν σταθερά για όλες τις παρατηρήσεις.

Ενώ η μέθοδος επιλογής χαρακτηριστικών δεν δημιουργεί νέα σύνολα χαρακτηριστικών εκπαίδευσης, αλλά αντίθετα συμπυκνώνει το προηγούμενο σύνολο χαρακτηριστικών, η μέθοδος εξαγωγής δεύτερης τάξης χαρακτηριστικών στοχεύει στη μείωση του πλεονασμού χαρακτηριστικών, ανάγοντας τις αρχικές διαστάσεις των εξαγόμενων χαρακτηριστικών σε ένα νέο χώρο χαμηλότερων διαστάσεων. Παραδείγματα ορισμένων αλγορίθμων εξαγωγής δεύτερης τάξης χαρακτηριστικών είναι η Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis), η Ανάλυση Ανεξάρτητων Συνιστωσών (Independent Components Analysis) και η Γραμμική Διακριτική Ανάλυση του Fisher (Fishers Linear Discriminant Analysis). Ο λόγος που γίνεται εφαρμογή αυτών των αλγορίθμων εκτός από τη μείωση διαστάσεων του δείγματος είναι και η βελτιστοποίηση των δεδομένων εκπαίδευσης σύμφωνα με ένα κριτήριο μέτρου ποιότητας των χαρακτηριστικών. Αντίστοιχα, ένα παράδειγμα αλγορίθμου επιλογής χαρακτηριστικών είναι ο Bag of Words (BoW), ο οποίος μειώνει το μέγεθος των δεδομένων εκπαίδευσης ομαδοποιώντας τα χαρακτηριστικά που είναι παρόμοια μεταξύ τους ως οπτικές λέξεις και μετατρέπει κάθε παρατήρηση ως ιστόγραμμα των παραγόμενων λέξεων. Στη συνέχεια ακολουθεί μια σύντομη επισκόπηση των περιγραφικών χαρακτηριστικών που χρησιμοποιήθηκαν στην παρούσα διπλωματική.

3.3 Στατικοί Περιγραφείς Χαρακτηριστικών

3.3.1 Περιγραφέας Τοπικών Δυναδικών Προτύπων - Local Binary Patterns

Τα Τοπικά Δυναδικά Πρότυπα περιεγράφηκαν πρώτη φορά από τον Harwood και στη συνέχεια παρουσιάστηκαν επισήμως ως μέθοδος ανάλυσης υφής το 1996 από τους Ojala. Έκτοτε χρησιμοποιούνται ευρύτατα σε διάφορες εφαρμογές καθώς αποτελούν ισχυρό και αποτελεσματικό εργαλείο για τον προσδιορισμό της υφής λόγω της διακριτικής τους ικανότητας και της χαμηλής υπολογιστικής πολυπλοκότητας. Ο LBP περιγραφέας αναπτύχθηκε αρχικά για την περιγραφή της υφής των εικόνων και βασίζεται στην ιδέα ότι δεν είναι απαραίτητο η εξέταση της εικόνας ως σύνολο αλλά αρκεί η εξέταση των τοπικών χαρακτηριστικών της, ο οποίος αποτυπώνει και κωδικοποιεί τοπικά χαρακτηριστικά εικόνων, δηλαδή την γειτονιά γύρω από κάθε εικονοστοιχείο, αφήνοντας αναλλοίωτες τις αποχρώσεις του γκρι. Ο πρότυπος τελεστής που αποδίδει μια τιμή σε κάθε εικονοστοιχείο της εικόνας έχει ως εξής:

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \quad (3.3.1)$$

,όπου p ο αριθμός των γειτόνων του κεντρικού εικονοστοιχείου και i_p, i_c οι τιμές του κεντρικού εικονοστοιχείου και των γειτονικών αντίστοιχα. Η συνάρτηση $s(x)$ ορίζεται ως:

$$s(x) = \begin{cases} 1 & \text{εάν } x \geq 0 \\ 0 & \text{εάν } x < 0 \end{cases} \quad (3.3.2)$$

Ένας βασικός περιορισμός του περιγραφέα LBP είναι ότι μικρές 3×3 γειτονίες δεν μπορούν να εξάγουν ικανοποιητικά αποτελέσματα σε εικόνες υψηλής κλίμακας. Η ανάγκη επίλυσης του περιορισμού αυτού, οδήγησε στην γενίκευση των LBP από τους Ojala το 2002, ώστε ο τελεστής να αποτελεί ένα απλό αλλά αποτελεσματικό τρόπο πολυ-διακριτικής ανάλυσης. Σύμφωνα με τη γενίκευση αυτή, ο καθορισμός των τιμών του LBP δε γίνεται σε μία γειτονία 3×3 αλλά στην περιφέρεια ενός κύκλου (P, R) όπου P το πλήθος των γειτονικών

εικονοστοιχείων που κατανέμονται ομοιόμορφα στην περιφέρεια του κύκλου και R η ακτίνα του κύκλου (Circular LBP). Παρακάτω παρουσιάζονται τρία χαρακτηριστικά παραδείγματα από γειτονιές διαφορετικής ακτίνας και πλήθους γειτόνων όπου οι συντεταγμένες προκύπτουν από τις σχέσεις:

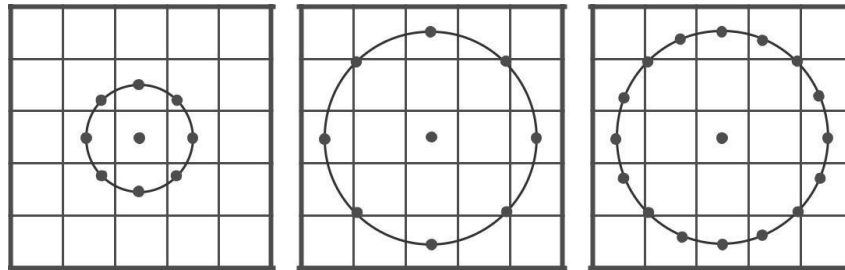
$$x_p = x_c + R * \cos\left(\frac{2\pi p}{P}\right) \quad (3.3.3)$$

$$y_p = y_c + R * \sin\left(\frac{2\pi p}{P}\right) \quad (3.3.4)$$

,όπου x_c, y_c οι συντεταγμένες του κεντρικού εικονοστοιχείου. Ο υπολογισμός των νέων τιμών των τοπικών χαρακτηριστικών LBP περιγράφεται από την παρακάτω σχέση:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) * 2^p \quad (3.3.5)$$

,όπου g_p, g_c οι τιμές των γειτονικών και του κεντρικού εικονοστοιχείου αντίστοιχα.



Εικόνα 3.1 Κυκλικές συμμετρικές γειτονιές για διαφορετικές τιμές P και R ((8,1),(8,2) και (16,2) αντίστοιχα).

Αξίζει να σημειωθεί πως ο υπολογισμός των $LBP_{P,R}$ γίνεται μόνο για τα εικονοστοιχεία που απέχουν R ή περισσότερο από τα άκρα της εικόνας, ενώ τα υπόλοιπα εικονοστοιχεία που δεν μπορούν να έχουν ένα πλήρες κύκλο R γύρω τους αμελούνται.

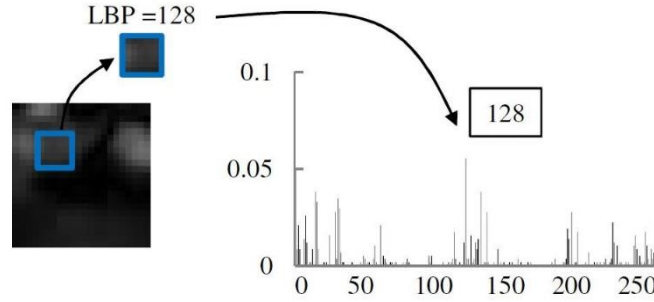
6	5	2
7	6	1
9	8	7

1	0	0
1		0
1	1	1

1	2	4
128		8
64	32	16

Εικόνα 3.2 Παράδειγμα υπολογισμού περιγραφέα $LBP_{8,1}$ με την αρχική εικόνα (αριστερά), το αποτέλεσμα του κατωφλίου (μέση) με μοτίβο=11110001 και τα βάρη από το κωδικοποιημένο άθροισμα (δεξιά), με αποτέλεσμα $LBP = 1 + 16 + 32 + 64 + 128 = 241$.

Ένα ακόμα σημαντικό μειονέκτημα των LBP όπως παρουσιάστηκαν μέχρι στιγμής είναι η μη ανθεκτικότητα τους στην περιστροφή της εικόνας ή στην σειρά με την οποία ορίζονται τα βάρη. Ο τρόπος που ορίζονται τα βάρη στις τιμές των εικονοστοιχείων δεν είναι πάντα ο ίδιος. Εάν γίνει η υπόθεση πως τα βάρη θα ορισθούν ξεκινώντας από το εικονοστοιχείο δεξιά από το κεντρικό και η εικόνα περιστρέφεται, οι τιμές των εικονοστοιχείων που εξετάζονται ως γειτονικά εικονοστοιχεία του κεντρικού θα μετακινηθούν αντίστοιχα, και συνεπακόλουθα θα επιστραφεί διαφορετική τιμή του τελεστή $LBP_{P,R}$. Διαφορετική τιμή του τελεστή θα προκύψει εάν αντίστροφα θεωρηθεί σταθερή η εικόνα και αλλάξει ο τρόπος που ορίζονται τα βάρη στα εικονοστοιχεία.



Εικόνα 3.3: Κατασκευή ιστογράμματος (8,1) LBP, όπου η τιμή του τελεστή LBP κάθε εικονοστοιχείου υπολογίζεται (αριστερά) και απεικονίζεται στο ιστόγραμμα (δεξιά).

Παρόλα αυτά η συγκεκριμένη αδυναμία, δεν ισχύει για πρότυπα που αποτελούνται αποκλειστικά από 1 ή αποκλειστικά από 0, καθώς όπως και αν γίνει η περιστροφή το τελικό αποτέλεσμα θα παραμένει ίδιο. Προκειμένου να απαλειφθεί το πρόβλημα που αναλύθηκε παραπάνω, της περιστροφής και της απόδοσης βαρών, οδήγησε στην περαιτέρω ανάπτυξη, του τελεστή LBP Uniform, όπου δίδεται μια μοναδική ταυτότητα με ένα δυαδικό πρότυπο ως εξής:

$$LBP_{P,R}^{u2} = \min\{ROR(LBP_{P,R,i}) | i = 0,1,2, \dots, P-1\} \quad (3.3.6)$$

,όπου η $ROR(x,i)$ πραγματοποιεί κυκλική ολίσθηση του δυαδικού αριθμού x προς τα δεξιά i φορές. Με τον τρόπο αυτό ομαδοποιούνται όλες οι εκδοχές περιστροφής ενός δυαδικού αριθμού σε μία και ταυτόχρονα ελαχιστοποιείται ο αριθμός των προτύπων που αντιστοιχούν σε συγκεκριμένα μικρό-χαρακτηριστικά της εικόνας. Σε μια γειτονιά εικονοστοιχείων $(8,R)$ τα $2^8 = 256$ διαφορετικά πρότυπα που προκύπτουν από τον απλό τελεστή LBP, μειώνονται σε 36 μοναδικά ανεξάρτητα περιστροφής πρότυπα με την εφαρμογή του LBP Uniform. Η μείωση του πλήθους των διαφορετικών προτύπων οδηγεί στη μείωση του μεγέθους του ιστογράμματος. Τα πρότυπα τα οποία αποτελούν θεμελιώδεις ιδιότητες της υψής και έχουν ως κοινό στοιχείο μια ομοιόμορφη κυκλική δομή ονομάστηκαν "ομοιόμορφα". Ομοιόμορφες δομές, λοιπόν, ονομάστηκαν εκείνες με τις λιγότερες μεταβάσεις από 0 σε 1 και αντίστροφα. Ο μαθηματικός ορισμός της ομοιομορφίας σε μία γειτονιά g δίνεται από τον παρακάτω τύπο:

$$U(G) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(s(g_{p-1} - g_c))| \quad (3.3.7)$$

Άρα με βάση τα παραπάνω, τα ομοιόμορφα ανεξάρτητα περιστροφής LBP ορίζονται ως:

$$LBP_{P,R}^{riu2}(x) = \begin{cases} \sum_{p=1}^{P-1} s(g_p - g_c) & \text{εάν } U(LBP_{P,R}) \geq 2 \\ P + 1 & \text{διαφορετικά} \end{cases} \quad (3.3.8)$$

Επομένως το σύνολο των ομοιόμορφων ανεξάρτητων περιστροφής προτύπων που μπορεί να εμφανιστούν σε μία γειτονιά $(8,P)$ παρουσιάζονται παρακάτω:



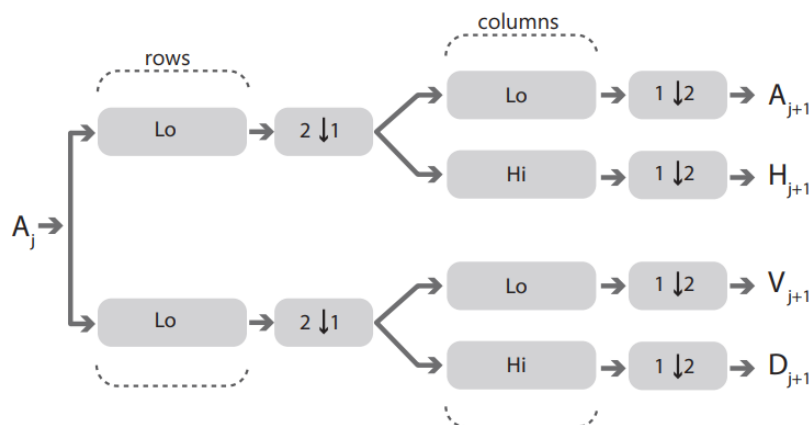
Εικόνα 3.4: Ομοιόμορφα ανεξάρτητα περιστροφής πρότυπα γειτονιάς $(8,P)$ όπου οι μαύρες βούλες αντιστοιχούν σε 0 και οι άσπρες σε 1.

3.3.2 Περιγραφείας Κυματιδίων - Wavelet

Όπως αναφέρθηκε και προηγουμένως με τον περιγραφέα δυαδικού προτύπου, η υφή μιας εικόνας μπορεί να αποτελέσει κομμάτι της περιγραφής του περιεχομένου της, ενώ η ανάλυση της είναι δυνατόν να οδηγήσει σε συμπεράσματα για τις παρατηρούμενες επιφάνειες και αντικείμενα. Ο μετασχηματισμός κυματιδίων (Wavelet) βρίσκει εφαρμογή σε πολλά διαφορετικά πεδία όπως στα μαθηματικά, τη φυσική, την ηλεκτρολογία, τη μηχανική όραση κ.α. Συνοπτικά ο μετασχηματισμός αυτός είναι μια αντιστρέψιμη, γρήγορη και γραμμική μέθοδος διαμερισμού του σήματος σε διαφορετικές υποζώνες συχνότητας και στη συνέχεια εξετάζει κάθε συστατικό σε σχέση με το μητρικό σήμα, το οποίο αντιπροσωπεύει τη βασική συνάρτηση αποσύνθεσης του κύματος – όμοια με τη ημιτονοειδή συνάρτηση στο μετασχηματισμό Fourier. Συνεπώς, ο μετασχηματισμός κύματος λειτουργεί ως φίλτρο ανάλυσης που εξετάζει ή λειτουργεί στα αποσυντιθέμενα κομμάτια αντί στο αρχικό σήμα. Προκειμένου να αναλυθεί το συστατικό κύμα, το μητρικό κλιμακώνεται, δηλαδή συμπιέζεται, διαστέλλεται, μετατοπίζεται κλπ. Ο συνεχής μετασχηματισμός κύματος (Continuous Wavelet Transform) μιας συνάρτησης $f(a,b)$ ορίζεται:

$$f(a,b) = \frac{1}{\sqrt{a}} \int f(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (3.3.9)$$

,όπου $\Psi(t)$ το μητρικό σήμα, a και b οι συντελεστές κλιμάκωσης και μετατόπισης αντίστοιχα, ενώ ο $\frac{1}{\sqrt{a}}$ λειτουργεί ως συντελεστής κανονικοποίησης. Ο Συνεχής Μετασχηματισμός Κυματιδίων αποτελεί μία από τις δύο μεγάλες κατηγορίες μετασχηματισμών κυματιδίων, ο δεύτερος είναι ο Διακριτός Μετασχηματισμός Κυματιδίων (Discrete Wavelet Transform). Ο τελευταίος βασίζεται στην κωδικοποίηση υποπεριοχών και παρέχει επαρκείς πληροφορίες τόσο για την ανάλυση όσο και για την ανακατασκευή του αρχικού σήματος, ενώ έχει διαπιστωθεί ότι ο χρόνος υπολογισμού και οι διαθέσιμοι πόροι που απαιτούνται είναι οι λιγότεροι. Στο μετασχηματισμό αυτό, το σήμα ακολουθεί μια διαδικασία στην οποία περνάει από υψηλής διέλευσης (high-pass) και χαμηλής διέλευσης (low-pass) φίλτρα με σκοπό την ανάλυση των υψηλών και των χαμηλών συχνοτήτων. Ο Διακριτός Μετασχηματισμός Κυματιδίων χρησιμοποιεί φίλτρα με διαφορετικές συχνότητες αποκοπής για να αναλυθεί το σήμα σε διαφορετικές κλίμακες, αποδίδοντας $3k+1$ υποζώνες, δηλαδή στο μητρικό σήμα (A_k) και σε τρεις συντελεστές τον ορθογώνιο, τον οριζόντιο και τον διαγώνιο (H_i, V_i, D_i , αντίστοιχα) για το επίπεδο αποσύνθεσης k και για κάθε υποεπίπεδο i , όπως παρουσιάζεται και παρακάτω:



Εικόνα 3.5: Αποσύνθεση ενός επιπέδου διακριτού σήματος από μία εικόνα με εφαρμογή φίλτρων. Τα Lo και Hi υποδηλώνουν τα φίλτρα υψηλής και χαμηλής διαπερατότητας ενώ τα $2 \downarrow 1$ και $1 \downarrow 2$ υποδηλώνουν την υποδειγματοληψία των στηλών και των γραμμών της εικόνας αντίστοιχα.

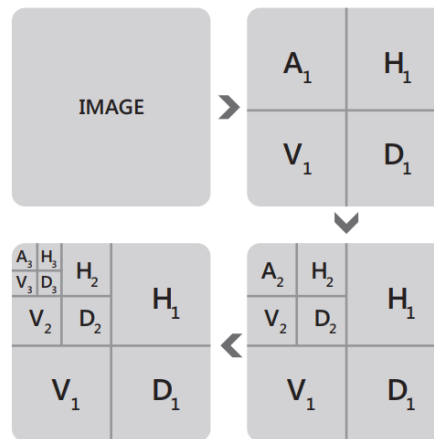
Λαμβάνοντας υπόψη μια εικόνα στη κλίμακα του γκρι ως ένα δισδιάστατο σήμα του οποίου οι συχνότητες αντιπροσωπεύουν τις μεταβολές της έντασης της εικόνας στις δύο κατευθύνσεις, σε γραμμές και στήλες, η αποσύνθεση των κυμάτων της στο επίπεδο k δίνεται:

$$[A_k, (H_1, V_1, D_1), \dots, (H_k, V_k, D_k)] \quad (3.3.10)$$

Συνεπώς, οι τρεις συντελεστές αποσύνθεσης (H_i, V_i, D_i) παρέχουν πληροφορίες σχετικά με το μέγεθος των ακμών της εικόνας που αναλύεται στις τρεις κατευθύνσεις στην ανάλυση του i -στού επιπέδου συχνότητας. Η συνολική ενέργεια μιας υποζώνης συχνότητας i μπορεί να υπολογισθεί τετραγωνίζοντας και αθροίζοντας τους αντίστοιχους συντελεστές:

$$E(I_i) = \sum_{x,y \in I} H_i(x,y)^2 + V_i(x,y)^2 + D_i(x,y)^2 \quad (3.3.11)$$

Η ιδέα εφαρμογής της χωρικής συχνότητας ως περιγραφέας χαρακτηριστικού μπορεί να βρεθεί σε έναν αριθμό αλγορίθμων για την ανίχνευση καπνού ή αερίων, ενώ έχουν δημοσιευθεί διάφορες προσεγγίσεις. Ο Wang πρότεινε μία μέθοδο που εκτελεί πλήρη αποσύνθεση κύματος μιας υποπεριοχής της εικόνας, η οποία χωρίς περαιτέρω επεξεργασία χρησιμοποιείται ως περιγραφέας χαρακτηριστικών για έναν ταξινομητή AdaBoost. Η μείωση των περιττών πληροφοριών σε αυτή την εφαρμογή επιτυγχάνεται με την δυνατότητα επιλογής χαρακτηριστικών που διαθέτει ο ταξινομητής Adaboost. Αποτελεί μία διαφορετική προσέγγιση, με χρήση περιγραφέα κυματιδίων, που εξάγει πληροφορίες από όλες τις υποζώνες αποσύνθεσης των κυμάτων σε όλα τα επίπεδα και στη συνέχεια υπολογίζει στατιστικά μέτρα, όπως μέση τιμή, τυπική απόκλιση κ.λπ. από την κατανομή των συντελεστών και την ένωση τους τη χρησιμοποιεί ως περιγραφέα για την εκπαίδευση ενός ταξινομητή.



Εικόνα 3.6: Αποσύνθεση τριών επιπέδων με χρήση κυματιδίων (Wavelets) μιας εικόνας.

3.3.3 Περιγραφέας Ιστογράμματος Προσανατολισμού Ακμών - Edge Orientation Histogram

Τα ιστογράμματα προσανατολισμού ακμών (Edge Orientation Histogram) εξάγουν πληροφορίες σχετικά με την υφή και το σχήμα μιας περιοχής μιας εικόνας, ενώ χρησιμοποιούνται ευρέως σε εφαρμογές αναγνώρισης αντικειμένων, όπως αναγνώριση προσώπου, ανίχνευση πεζών και ανίχνευση καπνού. Ο υπολογισμός ενός περιγραφέα ιστογράμματος προσανατολισμού ακμών είναι σχετικά απλός και ξεκινάει με τον υπολογισμό της κλίσης $G(x,y) = [f_x, f_y]$ σε κάθε (x,y) εικονοστοιχείο με κατάλληλη μέθοδο, όπως

παραδείγματος χάρη με τη μέθοδο Sobel. Το μέγεθος της ακμής και η κλίση της δίνονται από τους παρακάτω τύπους:

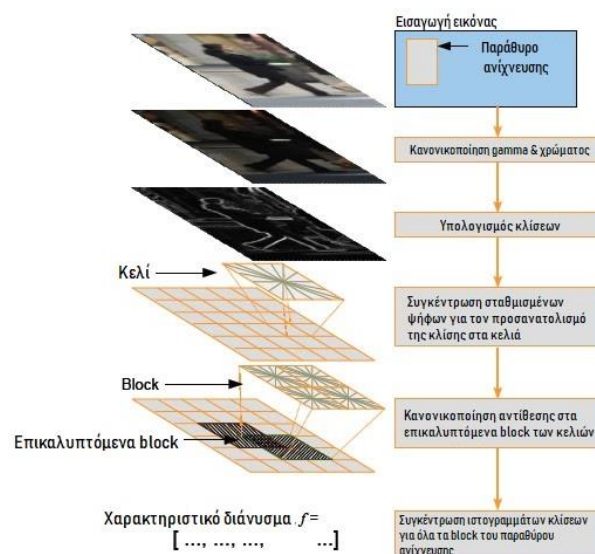
$$M(x, y) = \sqrt{f_x^2 + f_y^2} \quad (3.3.12)$$

$$\theta(x, y) = \arctan\left(\frac{f_y}{f_x}\right) \quad (3.3.13)$$

Στη συνέχεια κατασκευάζεται το ιστόγραμμα προσανατολισμού ακμών στο διάστημα $[0^\circ - 180^\circ]$. Το βήμα K του ιστογράμματος ορίζεται από τον χρήστη, όπου μικρό K δεν δίνει καλά διακριτά αποτελέσματα ενώ αντίστοιχα μεγάλο K δεν καταφέρνει να αποδώσει χαρακτηριστικά υπό περιοχών. Το ιστόγραμμα κλίσεων διαβαθμίσεων που αναλύεται παρακάτω, είναι ουσιαστικά μια επέκταση του περιγραφέα ιστογράμματος προσανατολισμού ακμών και αποτελεί μία από τις πιο διαδεδομένες μεθόδους ανίχνευσης ανθρώπων. Ενώ ο περιγραφέας ιστογράμματος προσανατολισμού ακμών περιγράφει την γενική υφή ολόκληρης της εικόνας, ο περιγραφέας ιστογράμματος κλίσεων διαβαθμίσεων είναι σχεδιασμένος για να συλλαμβάνει ορισμένες δομές εξάγοντας και συλλέγοντας ιστογράμματα προσανατολισμού ακμών από ένα παράθυρο ανίχνευσης προκαθορισμένου μεγέθους.

3.3.4 Περιγραφέας Ιστογράμματος Κλίσεων Διαβαθμίσεων - Histogram of Oriented Features

Η μέθοδος ιστογράμματος κλίσεων διαβαθμίσεων ή βάθμωσης είναι μια μέθοδος 'πυκνής' εξαγωγής χαρακτηριστικών από εικόνες, ένας δηλαδή περιγραφέας χαρακτηριστικών. Ένας περιγραφέας, όπως αναφέρθηκε και παραπάνω, έχει ως σκοπό να περιγράψει ένα αντικείμενο με τέτοιο τρόπο ώστε το ίδιο το αντικείμενο να μπορεί να παράγει όσον το δυνατόν πλησιέστερα για τον ίδιο περιγραφέα χαρακτηριστικών όταν προβληθεί υπό διαφορετικές συνθήκες, γεγονός που καθιστά ευκολότερη την ταξινόμηση. Με τον όρο "πυκνή", εννοείται ότι εξάγει χαρακτηριστικά για όλες τις θέσεις της εικόνας και όχι μόνο για μία τοπική γειτονιά σημείων κλειδιών.



Εικόνα 3.7: Διαδικασία κατασκευής ιστογράμματος του περιγραφέα βάθμωσης

Ο στόχος του ιστογράμματος κλίσεων είναι η καταγραφή των ορίων μιας περιοχής με τη συλλογή πληροφοριών σχετικά με τις κλίσεις στην εικόνα και ταυτόχρονα η δημιουργία ενός ιστογράμματος από τον προσανατολισμό αυτής της διαβάθμισης. Το ιστόγραμμα έχει 36 διακριτές τιμές που καλύπτουν το σύνολο των 360° με βήμα 10° . Ένας ανιχνευτής κλίσεων συνήθως χρησιμοποιεί ένα παράθυρο ανίχνευσης με πλάτος 64 και ύψος 128 εικονοστοιχείων. Ο υπολογισμός του περιγραφέα συχνά γίνεται σε κελιά των 8 επί 8 εικονοστοιχείων εντός του παραθύρου ανίχνευσης, όπου τα κελιά αυτά οργανώνονται σε επικαλυπτόμενα μπλοκ. Σε κάθε κελί υπολογίζονται οι οριζόντιες και κατακόρυφες κλίσεις για κάθε εικονοστοιχείο.

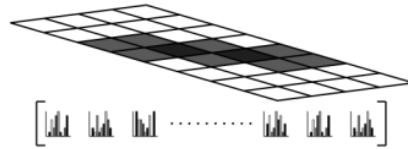
$$G_x(y, x) = Y(y, x + 1) - Y(y, x - 1) \quad (3.3.14)$$

$$G_y(y, x) = Y(y + 1, x) - Y(y - 1, x) \quad (3.3.15)$$

όπου $Y(y, x)$ είναι η ένταση των εικονοστοιχείων στις συντεταγμένες x, y και G_x και G_y η οριζόντια και η κάθετη κλίση αντίστοιχα. Το μέγεθος και η γωνία της κλίσης υπολογίζονται:

$$G(y, x) = \sqrt{G_x(y, x)^2 + G_y(y, x)^2} \text{ και } \theta = \arctan\left(\frac{G_y(y, x)}{G_x(y, x)}\right) \quad (3.3.16)$$

Συνεπώς, από ένα κελί 8×8 εικονοστοιχείων, 64 διανύσματα κλίσης παράγονται και τοποθετούνται σε ένα ιστόγραμμα 9 βημάτων (bins). Το ιστόγραμμα κυμαίνεται από 0 έως 180 μοίρες, με 20 μοίρες ανά βήμα. Για κάθε διαβάθμιση κλίσης, η συμβολή της στο ιστόγραμμα δίνεται από το μέγεθος του διανύσματος, καθώς μεγάλες κλίσεις έχουν μεγαλύτερη επίδραση στο ιστόγραμμα.



Εικόνα 3.8: Η συνένωση ιστογραμμάτων επικαλυπτόμενων block.

Στη περίπτωση που κάποια κλίση δεν ταυτίζεται με κάποιο από τα βήματα του ιστογράμματος, τότε διαχωρίζεται με τη συμβολή της κλίσης στα βήματα που βρίσκεται ανάμεσα. Για παράδειγμα, εάν κάποια κλίση έχει γωνία 85° τότε προστίθεται το $\frac{1}{4}$ του μεγέθους στο βήμα 70° και τα $\frac{3}{4}$ του μεγέθους στο βήμα 90° . Καθώς διαφορετικές εικόνες μπορεί να έχουν διαφορετική αντίθεση, γίνεται κανονικοποίηση των ιστογραμμάτων έτσι ώστε να μην επηρεάζονται από αλλαγές έντασης, προκειμένου ο δημιουργούμενος δείκτης να είναι ανεξάρτητος από τις παραλλαγές φωτισμού. Οι δημοφιλέστεροι τύποι κανονικοποίησης είναι:

$$L1norm = \frac{v}{\sqrt{|v|_2^2 + s^2}}, L2norm = \frac{v}{|v|_1 + s^2}, L1sqnorm = \sqrt{\frac{v}{|v|_1 + s}} \quad (3.3.17)$$

3.4 Δυναμικοί Περιγραφείς Χαρακτηριστικών

3.4.1 Περιγραφέας Ιστογράμματος Οπτικής Ροής - Histogram of Optical Flow

Ως οπτική ροή ορίζεται η μέθοδος υπολογισμού της σχετικής κίνησης των εικονοστοιχείων μίας εικόνας του βίντεο σε σχέση με την προηγούμενη. Ο υπολογισμός της ροής μετέπειτα

αποθηκεύεται ως διάνυσμα που περιγράφει την μετατόπιση κάθε εικονοστοιχείου σε κάθε εικόνα του βίντεο και χρησιμοποιείται για την απεικόνιση της κίνησης. Υποθέτοντας ότι ένα εικονοστοιχείο (x,y) σε μία εικόνα τη χρονική στιγμή t του βίντεο έχει τιμή $I(x,y,t)$. Το εικονοστοιχείο διανύει απόσταση (dx,dy) στην επόμενη εικόνα του βίντεο σε χρονικό διάστημα dt . Υποθέτοντας ξανά πως αυτά τα εικονοστοιχεία είναι ίδια και η τιμή της έντασης τους δεν έχει μεταβληθεί τότε:

$$I(x,y,t) = I(x + dx, y + dy, t + dt) \quad (3.4.1)$$



Εικόνα 3.9: Διαδικασία υπολογισμού περιγραφέα οπτικής ροής.

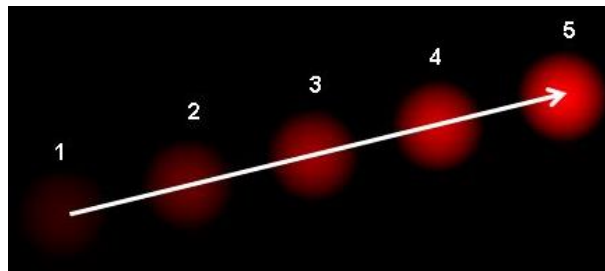
Η διαδοχική εικόνα για τη χρονική στιγμή $t+1$ του βίντεο περιγράφεται από τη μη γραμμική συνάρτηση οπτικής ροής:

$$I(x,t) = I(x + u, y + v, t + 1) \quad (3.4.2)$$

,όπου (u, v) το διάνυσμα της οπτικής ροής, δηλαδή η μετατόπιση του εικονοστοιχείου (x, y) από την εικόνα τη χρονική στιγμή t του βίντεο στην εικόνα τη χρονική στιγμή $t+1$ του βίντεο. Εφαρμόζοντας το θεώρημα της γραμμικοποίησης, η συνάρτηση μετατρέπεται:

$$I_x(x,t)u + I_y(y,t)v + I_t(x,t) = 0 \quad (3.4.3)$$

,όπου I_x, I_y οι χωρικές παράγωγοι και I_t η χρονική παράγωγος αντιστοίχα.



Εικόνα 3.10: Το διάνυσμα οπτικής ροής ενός κινούμενου αντικειμένου σε μια ακολουθία στιγμιότυπων βίντεο.

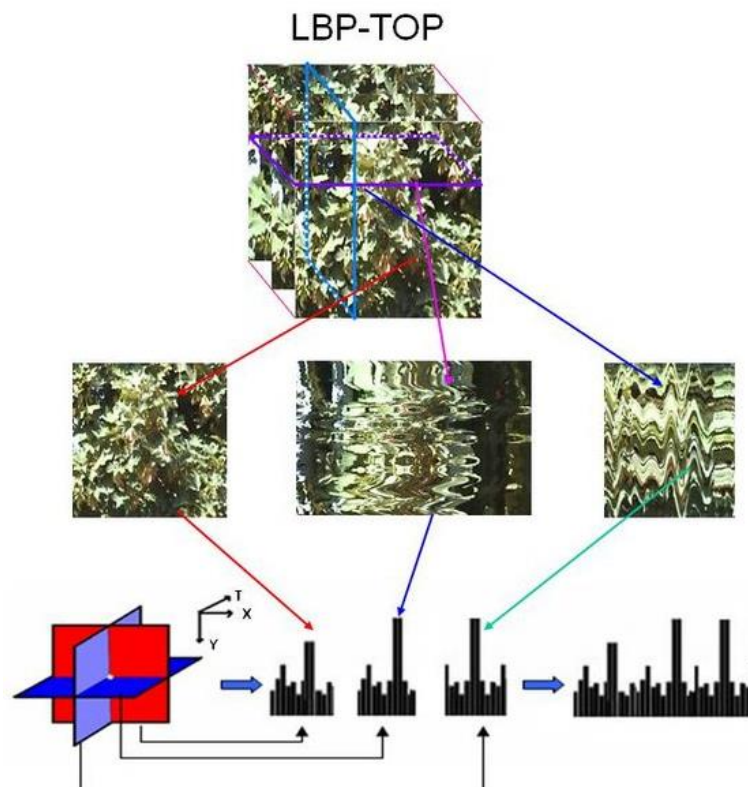
Για την επίλυση της παραπάνω εξίσωσης υπάρχουν δύο γενικές προσεγγίσεις, η μέθοδος Lucas-Kanade και η μέθοδος Horn-Schunck. Η πρώτη μέθοδος, διευρύνει το πρόβλημα στην γειτονιά του εικονοστοιχείου, υποθέτοντας ότι η ταχύτητα όλων των εικονοστοιχείων μέσα σε αυτήν την γειτονιά είναι σταθερή και υπολογίζει ένα 'τοπικό' διάνυσμα οπτικής ροής. Η δεύτερη μέθοδος εισάγει μια καινούρια συνάρτηση για την επίλυση της εξίσωσης της οπτικής ροής για όλη την εικόνα εισάγοντας μια μεταβλητή εξομάλυνσης α :

$$E(u, v) = \int (\nabla I(x, t)u + \nabla I(y, t)v + I_t(x, t))^2 + \alpha(|\nabla u|^2 + |\nabla v|^2) dx dy = 0 \quad (3.4.4)$$

Για την κατασκευή του ιστογράμματος της οπτικής ροής (HOF), συλλέγονται τα διανύσματα της οπτικής ροής για W εικόνες του βίντεο, ταξινομούνται στο διάστημα $(0^\circ, 180^\circ)$ και τοποθετούνται σε ένα ιστόγραμμα K βημάτων (bins) σύμφωνα με τη γωνία της κλίσης. Η παράμετροι W (αριθμός εικόνων βίντεο) και K (bins) ορίζονται από το χρήστη. Στη συνέχεια το ιστόγραμμα κανονικοποιείται προκειμένου να επιτευχθεί ανεξαρτησία στην κλίμακα.

3.4.2 Περιγραφέας Τοπικών Δυναδικών Προτύπων Τριών Διαστάσεων - Local Binary Patterns from Three Orthogonal Planes

Ένας ακόμη χωροχρονικός περιγραφέας χαρακτηριστικών που συνδυάζει δυναμικά την υφή είναι η επέκταση του προαναφερθέντος περιγραφέα τοπικών δυναδικών προτύπων στις τρεις διαστάσεις. Καθώς ο διδιάστατος περιγραφέας των τοπικών δυναδικών προτύπων (LBP) είναι ένας στατικός περιγραφέας χαρακτηριστικών δεν μπορεί να λάβει υπόψιν την μορφολογία της υφής, δηλαδή πως μετασχηματίζεται χρονικά, στα στιγμιότυπα ενός βίντεο. Οι Zhao και Pietikäinen το 2007 πρότειναν μία παραλλαγή του τοπικού περιγραφέα σε χωροχρονικό περιγραφέα τοπικών δυναδικών προτύπων βάθους (Volume LBP), με σκοπό να χαρακτηριστούν δυναμικά οι υφές στις ακολουθίες των βίντεο. Ο δυναμικός περιγραφέας τριών διαστάσεων, αποδίδει την δυναμική υφή σε μία ακολουθία εικόνων W , όπου W ο αριθμός των στιγμιότυπων του βίντεο που προσδιορίζεται από τον χρήστη, κατασκευάζοντας έναν τρισδιάστατο κύβο τριών αξόνων. Οι άξονες X, Y υποδηλώνουν τους χωρικούς άξονες των εικόνων του βίντεο, ενώ ο άξονας T είναι ο άξονας του χρόνου, δηλαδή ο αριθμός των στιγμιότυπων για τον οποίο υπολογίζεται ο συγκεκριμένος περιγραφέας.



Εικόνα 3.11: Ο περιγραφέας τοπικών δυναδικών προτύπων τριών διαστάσεων

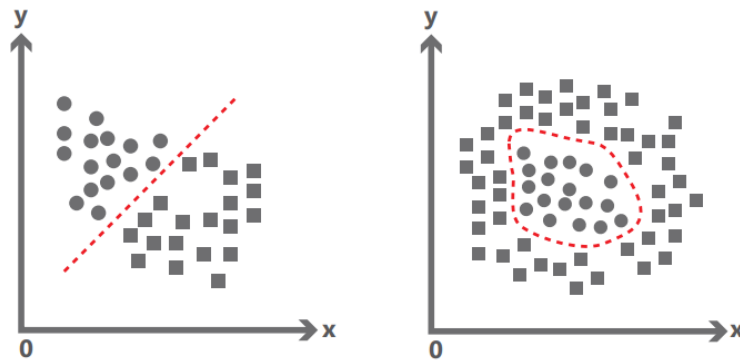
Με αυτόν τον τρόπο, η γειτονιά του κάθε εικονοστοιχείου ορίζεται έτσι σε ένα τρισδιάστατο χώρο. Στη συνέχεια υπολογίζεται το ανάγλυφο της περιστροφής του κάθε εικονοστοιχείου και ανάλογα με την επιλογή του κατάλληλου περιγραφέα LBP (απλό, circular, uniform, κ.α.) κατασκευάζεται το ιστόγραμμα χαρακτηριστικών. Συνεπώς, ο VLBP περιγραφέας συνδυάζει την κίνηση και την εμφάνιση μαζί για να περιγράψει δυναμικά την υφή. Προκειμένου η κατασκευή του VLBP να καταστεί υπολογιστικά εύκολη και απλή, εισήχθη ένας επιπλέον περιγραφέας που αποτελεί ανάπτυξη του προηγούμενου, και βασίστηκε στις συσχετίσεις των τοπικών δυαδικών προτύπων σε τρία ορθογώνια επίπεδα (LBP-Three Orthogonal Planes). Ο περιγραφέας LBPTOP εξετάζει τρία ορθογώνια επίπεδα: XY, XT και YT και συνενώνει τα τοπικά στατιστικά στοιχεία από αυτές τις τρεις κατευθύνσεις. Πιο αναλυτικά, το επίπεδο XY αντιπροσωπεύει πληροφορίες εικόνας, το επίπεδο XT δίνει μια οπτική εντύπωση της μεταβολής μίας σειράς στο χρόνο, ενώ αντίστοιχα, το επίπεδο YT αντιπροσωπεύει την μεταβολή μίας στήλης στο χρόνο. Οι τελεστές LBP που υπολογίζονται από όλα τα εικονοστοιχεία των τριών επιπέδων συμβολίζονται ως XY-LBP, XT-LBP και YT-LBP αντίστοιχα, και τα ιστογράμματα που κατασκευάζονται από αυτά τα τρία επίπεδα ενώνονται σε ένα ιστόγραμμα του οποίου η διάσταση εξαρτάται από το είδος του στατικού περιγραφέα LBP που θα χρησιμοποιηθεί. Εάν χρησιμοποιηθεί για παράδειγμα ένας τελεστής LBP Uniform σε μία γειτονιά (8,1), το τελικό ιστόγραμμα του περιγραφέα LBP-TOP θα έχει μήκος $3 \times 59 = 177$.

$$\text{LBPTOP}_{P_{XY}P_{XT}P_{YT}R_{XY}R_{XT}R_{YT}} = [\text{LBP}_{P,R}(XY)\text{LBP}_{P,R}(XT)\text{LBP}_{P,R}(YT)] \quad (3.4.5)$$

Ο συγκεκριμένος περιγραφέας τοπικών δυαδικών προτύπων τριών διαστάσεων έχει χρησιμοποιηθεί σε αρκετές εφαρμογές αναγνώρισης έκφρασης προσώπου και αναγνώρισης ανθρώπων πραξέων.

3.5 Αλγόριθμοι ταξινόμησης

Ο όρος ταξινομητής ή αλγόριθμος ταξινόμησης σηματοδοτεί ένα μηχανισμό που κατατάσσει μια τιμή σε μία από τις διάφορες κλάσεις. Συνοπτικά, ένας ανιχνευτής ενός αντικειμένου απαρτίζεται από δύο κύρια στοιχεία. Το πρώτο είναι ένα είδος περιγραφέα, ο οποίος μπορεί να αποτελείται από συνδυασμό περισσότερων περιγραφέων και σκοπός του είναι να μπορεί να περιγράψει ικανοποιητικά το αντικείμενο ως προς τα χαρακτηριστικά του. Το δεύτερο στοιχείο είναι ένας ταξινομητής ο οποίος εκπαιδεύεται από τον περιγραφέα του αντικειμένου προκειμένου να μπορεί να διαχωρίσει την εικόνα από το αντικείμενο ενδιαφέροντος και των αντικειμένων υποβάθρου, ταξινομώντας τα ως θετικά ή αρνητικά αντίστοιχα. Η γενική διαδικασία που ακολουθείται κατά την επιβλεπόμενη ταξινόμηση ξεκινάει με την επιλογή της συνάρτησης απόφασης του ταξινομητή. Στη συνέχεια εξάγονται τόσο από τα θετικά όσο και από τα αρνητικά δείγματα εξάγεται ο επιθυμητός περιγραφέας και με τα αποτελέσματα εκπαιδεύεται ο ταξινομητής και συνεπώς μια συνάρτηση απόφασης, με την οποία αποφασίζει ο ταξινομητής την ύπαρξη ή μη του επιθυμητού αντικειμένου στην εικόνα. Όπως είναι κατανοητό, για την σωστή κατασκευή ενός ανιχνευτή αντικειμένου, απαιτείται η ορθή επιλογή των κατάλληλων περιγραφέων και του κατάλληλου μηχανισμού ταξινόμησης. Στην περίπτωση της δυαδικής ταξινόμησης, μία τιμή μπορεί να ανήκει σε μία από τις δύο κλάσεις, όπου κάθε μία χαρακτηρίζει τις "θετικές" και τις "αρνητικές" τιμές αντίστοιχα, ή ειδικότερα στην περίπτωση ανίχνευσης αερίου, σε "αέριο" ή "μη αέριο". Η τιμή που ταξινομείται από έναν αλγόριθμο ταξινόμησης μπορεί να περιγράφεται από ένα n-οστό διάνυσμα σε ένα n-διάστατο χώρο. Ένα απλό παράδειγμα δυαδικής ταξινόμησης είναι για παράδειγμα η ταξινόμηση ενός συνόλου θερμοκρασιακών τιμών σε "θερμές" ή "ψυχρές".



Εικόνα 3.12: Γραμμικά διαχωρίσιμα (αριστερά) και μη γραμμικά διαχωρίσιμα (δεξιά) δεδομένα δυοδιάστατου χώρου, συμπεριλαμβανομένων των συναρτήσεων απόφασης αντίστοιχα.

Σε μια τέτοια περίπτωση, αρκεί να έχει προηγουμένως τεθεί ένα προκαθορισμένο όριο που λειτουργεί ως τιμή απόφασης για τον αλγόριθμο στον μονοδιάστατο χώρο και το οποίο διαχωρίζει σαφώς τις τιμές στις δύο πιθανές κλάσεις. Αυτή η περίπτωση αποτελεί τον απλούστερο μηχανισμό ταξινόμησης τιμών, καθώς αποτελεί ένα μονοδιάστατο μηχανισμό ταξινόμησης με κατώφλι. Συνήθως, τα περισσότερα προβλήματα ταξινόμησης είναι πιο περίπλοκα και σχεδόν αδύνατο να συμπεκνωθούν σε μονοδιάστατες τιμές. Συνεπώς η επίλυση τέτοιων περίπλοκων προβλημάτων ταξινόμησης υλοποιείται με τη χρήση αλγορίθμων μηχανικής εκμάθησης. Όσο μεγαλύτερος ο βαθμός των διαστάσεων της τιμής που κατηγοριοποιείται (όπως για παράδειγμα ένας περιγραφέας ιστογράμματος έντασης μιας μονόχρωμης εικόνας 256 διαστάσεων που χρησιμοποιείται ως διάνυσμα χαρακτηριστικών για την κατηγοριοποίηση εικόνων) και όσο πιο περίπλοκη η μέθοδος ταξινόμησης, τόσο πιο δύσκολο γίνεται να βρεθεί εμπειρικά ένα κατώφλι που χρησιμοποιείται για τον διαχωρισμό των τιμών στις διαφορετικές κλάσεις. Υπάρχουν δύο στοιχειώδη είδη αλγορίθμων μηχανικής μάθησης, που διαφοροποιούνται ως προς τον τρόπο που με τον οποίο δημιουργείται μία συνάρτηση απόφασης.

Ο πρώτος τρόπος ονομάζεται Επιβλεπόμενη Μάθηση (Supervised Learning), όπου στόχος της διαδικασίας αυτής είναι ο χαρακτηρισμός δεδομένων με βάση ορισμένα δεδομένα εκπαίδευσης [14]. Τα δεδομένα εκπαίδευσης είναι ένα σύνολο θετικών και αρνητικών δεδομένων που χρησιμοποιούνται για την εκπαίδευση των μοντέλων. Συνεπώς, στην επιβλεπόμενη μάθηση κάθε συνάρτηση αποτελείται από ένα σύνολο εισόδου (συνήθως ένα διάνυσμα από χαρακτηριστικά) και μία τιμή εξόδου. Ο απώτερος σκοπός των αλγορίθμων επιβλεπόμενης μάθησης είναι η δημιουργία ενός μοντέλου το οποίο μπορεί να γενικεύσει τη συνάρτηση αυτή και να χαρακτηρίσει διαφορετικά σύνολα δεδομένων από τα δεδομένα εκπαίδευσης.

Ο δεύτερος τρόπος ονομάζεται Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning), όπου στόχος της είναι η ανακάλυψη πιθανής δομής των δεδομένων που μπορεί να παρουσιάζουν μη χαρακτηρισμένα δεδομένα. Δηλαδή, ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Γνωστά παραδείγματα μεθόδων Μη Επιβλεπόμενης Μάθησης αποτελούν οι τεχνικές ομαδοποίησης (clustering), όπως ο γνωστός αλγόριθμος k-means, που έχει ως στόχο την κατανομή των παρατηρήσεων σε ομάδες έτσι ώστε κάθε παρατήρηση να ανήκει στην ομάδα με τον κοντινότερο μέσο, ο οποίος αποτελεί και το χαρακτηριστικό δείγμα της ομάδας. Καθώς στην παρούσα διπλωματική, οι μέθοδοι ταξινόμησης που χρησιμοποιήθηκαν εμπίπτουν στην κατηγορία της επιβλεπόμενης μάθησης, η μη επιβλεπόμενη μάθηση δεν αναλύεται περαιτέρω. Οι αλγόριθμοι της επιβλεπόμενης μάθησης υπάρχουν σε διάφορες παραλλαγές, καθένας με τα δικά του πλεονεκτήματα και μειονεκτήματα και δικό του πεδίο εφαρμογής. Από τους αλγορίθμους επιβλεπόμενης μάθησης ξεχωρίζουν οι Ταξινομητές Bayes (Naïve Bayes Classifiers), τα Νευρωνικά Δίκτυα (Artificial Neural Networks), οι Μηχανές Διανυσμάτων

Υποστήριξης (Support Vector Machines) καθώς και οι τεχνικές προσαρμοστικής "ώθησης" (AdaBoost). Η χρήση της τεχνικής των νευρωνικών δικτύων, που είναι και η πιο διαδεδομένη λόγω των πλεονεκτημάτων που παρουσιάζει, απαιτεί την παροχή μεγάλου αριθμού εκπαίδευσης στοιχείο που δεν ήταν διαθέσιμο στην παρούσα διπλωματική και συνεπώς η ταξινόμηση των δεδομένων επιτεύχθηκε με χρήση των ταξινομητών SVM και AdaBoost, οι οποίοι αναλύονται παρακάτω.

3.5.1 Μηχανές Διανυσμάτων Υποστήριξης - Support Vector Machines

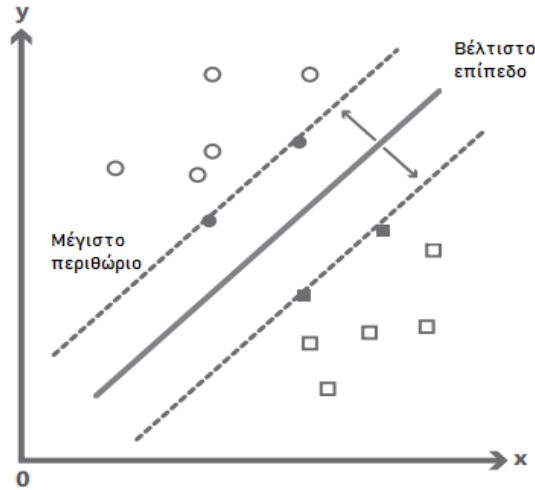
Στην παράγραφο που ακολουθεί, παρουσιάζεται μια σύντομη εισαγωγή στις θεμελιώδεις έννοιες των μηχανών διανυσμάτων υποστήριξης, ενώ περισσότερες λεπτομέρειες και μεγαλύτερη ανάλυση του μαθηματικού υποβάθρου μπορούν να βρεθούν στις αναφορές [15], [16], [17]. Ο αλγόριθμος μηχανής διανυσμάτων υποστήριξης αποτελεί έναν από τους πιο δημοφιλείς αλγορίθμους επιβλεπόμενης μάθησης. Η βασική ιδέα αυτής της μεθόδου ταξινόμησης είναι η κατηγοριοποίηση των δεδομένων με την εύρεση ενός βέλτιστου υπέρ-επιπέδου που διαχωρίζει τα δεδομένα με το μέγιστο δυνατό περιθώριο. Στην περίπτωση που τα δεδομένα δεν μπορούν να διαχωριστούν γραμμικά, με τη χρήση κατάλληλων απεικονίσεων μεταφέρονται τα δεδομένα μεταφέρονται σε μεγαλύτερη διάσταση προκειμένου να επιτευχθεί ο διαχωρισμός τους. Τα όρια των κλάσεων που κατηγοριοποιούνται τα δεδομένα καθορίζονται από τα περισσότερο εξωτερικά διανύσματα, που ονομάζονται διανύσματα υποστήριξης. Δεδομένου ενός n -διάστατου δείγματος x_i , όπου $i = 1, \dots, N$, και N το πλήθος των γραμμικά διαχωριζόμενων τιμών, και $y_i = +1$ και $y_i = -1$ οι τιμές που κατηγοριοποιούν τα δεδομένα στις δύο κλάσεις, ως θετικά ή αρνητικά, ο τρόπος διαχωρισμούς περιγράφεται από τον παρακάτω μαθηματικό τύπο:

$$y_i(w * x_i + b) \geq 1, \quad i = 1, \dots, N \quad (3.5.1)$$

Το βέλτιστο επίπεδο διαχωρισμού των κατηγοριών περιγράφεται ως:

$$D(x) = w * x_i + b = c, \quad -1 < c < 1 \quad (3.5.2)$$

Το w και το b υποδηλώνουν τον προσανατολισμό και την μετατόπιση του επιπέδου από την αρχή, αντίστοιχα. Καθώς ο αριθμός των επιπέδων που μπορούν να διαχωρίσουν δύο κλάσεις είναι άπειρος, ενώ ταυτόχρονα ζητούμενο του αλγορίθμου είναι η επίτευξη της βέλτιστης απόδοσης για σύνολα δεδομένων με παρόμοιες κατανομές τιμών, το επίπεδο διαχωρισμού τοποθετείται όσο το δυνατόν πιο μακριά και από τα δύο σύνολα δεδομένων, δεδομένου ότι η θέση αυτή επιτρέπει τη σωστή ταξινόμηση ακόμα και αν οι τιμές των δεδομένων διαφέρουν σε κάποιο βαθμό. Η απόσταση μεταξύ δύο ομάδων ονομάζεται όριο δ του ταξινομητή.



Εικόνα 3.13: Οπτικοποίηση ταξινόμησης με διανύσματα υποστήριξης ενός γραμμικά διαχωρίσιμου δυαδικού συνόλου. Τα τετράγωνα και οι κύκλοι στο γράφημα αντιπροσωπεύουν τις δύο κλάσεις +1 και -1 αντίστοιχα.

Το μέγιστο όριο δ επιτυγχάνεται όταν η συνάρτηση του επιπέδου γίνεται μηδέν $D(x) = 0$. Δεδομένου επίσης της ευκλείδειας απόστασης $|D(x)|/\|w\|$ μεταξύ μιας τιμής x και του επιπέδου, το βέλτιστο επίπεδο διαχωρισμού υπολογίζεται αν όλες οι τιμές πληρούν το παρακάτω κριτήριο:

$$\frac{y_i D(x_k)}{w} \geq \delta, \quad i = 1, \dots, N \quad (3.5.3)$$

Υποθέτοντας ότι (w, b) είναι μία λύση, όλες οι πολλαπλάσιες λύσεις (aw, ab) αποτελούν λύσεις και αυτές, συνεπώς ο περιορισμός $\delta\|w\| = 1$ εξαλείφεται. Αυτές οι δύο εξισώσεις μπορούν να χρησιμοποιηθούν για να βρεθεί η ελάχιστη τιμή w που ικανοποιεί την εξίσωση των συνόρων της κλάσης των δεδομένων:

$$y_i(w * x_i + b), \quad i = 1, \dots, N \quad (3.5.4)$$

Το τετραγωνικό πρόβλημα βελτιστοποίησης περιγράφεται ως εξής:

$$\text{ελαχιστοποίηση: } Q(w, b) = \frac{1}{2} \|w\|^2 \quad (3.5.5)$$

$$\text{περιορισμοί: } y_i(w * x_i + b), \quad i = 1, \dots, N \quad (3.5.6)$$

Η τετραγωνική συνάρτηση λύνεται με την εισαγωγή πολλαπλασιαστών Lagrange a_i (με $i = 1, \dots, N$) στη συνάρτηση ελαχιστοποίησης:

$$Q(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N a_i [y_i(x_i * w + b) - 1]. \quad (3.5.7)$$

Η παραπάνω εξίσωση Lagrange, για να ελαχιστοποιηθεί σύμφωνα με τους περιορισμούς για τα w και b , πρέπει να ικανοποιούνται οι παρακάτω συνθήκες:

$$\frac{\partial Q(w, b, a)}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N a_i y_i x_i \quad (3.5.8)$$

$$\frac{\partial Q(w, b, a)}{\partial b} = 0 \Rightarrow \sum_{i=1}^N a_i y_i = 0 \quad (3.5.9)$$

Συνδυάζοντας τις παραπάνω εξισώσεις το πρόβλημα μοντελοποιείται ως εξής:

$$\text{μεγιστοποίηση: } Q(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j x_i x_j \quad (3.5.10)$$

$$\text{περιορισμοί: } a_i \geq \forall i, \quad \sum_{i=1}^N a_i y_i = 0 \quad (3.5.11)$$

Η μεγιστοποίηση μπορεί να επιτευχθεί με οποιοδήποτε αλγόριθμο επίλυσης τετραγωνικού προβλήματος. Από το σύνολο των δεδομένων κάθε τιμή x_i που σχετίζεται με μία μη μηδενική τιμή a_i ανήκει στην ομάδα των διανυσμάτων υποστήριξης καθώς το διανυσματικό υπέρ-επίπεδο καθορίζεται αποκλειστικά από αυτά τα διανύσματα. Οι τιμές αυτές x_i συμβολίζονται ως x_s , $s \in S$, όπου το S συμβολίζει το σύνολο των διανυσμάτων υποστήριξης. Ένα παράδειγμα διανύσματος υποστήριξης που ικανοποιεί και την παραπάνω εξίσωση είναι:

$$y_s(w * x_s + b) = 1 \quad (3.5.12)$$

,όπου αν αντικατασταθεί το $w = \sum_{i=1}^N a_i y_i x_i$, υπολογίζεται η συνάρτηση απόφασης αυστηρών ορίων:

$$D(x) = b + \sum_{i \in S} a_i y_i x_i * x \quad (3.5.13)$$

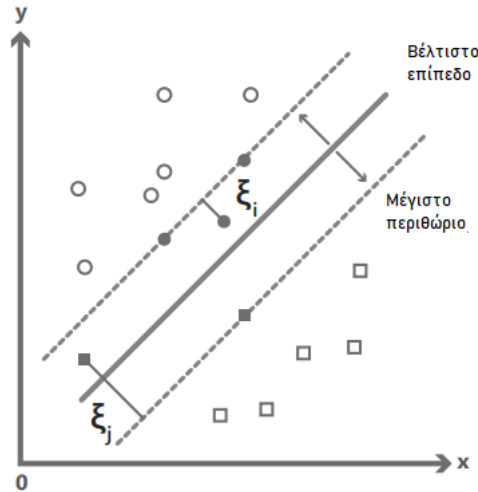
,η οποία επιτρέπει την ταξινόμηση μιας νέας άγνωστης τιμής x από ένα γραμμικά διαχωριζόμενο σύνολο δεδομένων ως:

$$\begin{cases} \text{Κατηγορία} + 1 & \text{if } D(x) > 0 \\ \text{Κατηγορία} - 1 & \text{if } D(x) < 0 \end{cases} \quad (3.5.14)$$

Παρόλα αυτά, η πλειονότητα των συνόλων δεδομένων δεν είναι γραμμικά διαχωρίσιμα, με αποτέλεσμα ένας ταξινομητής διανυσμάτων υποστήριξης αυστηρών ορίων να μη μπορεί να λύσει το πρόβλημα διαχωρισμού, αφού ένα τέτοιο επίπεδο διαχωρισμού μπορεί να μην υπάρχει. Προκειμένου να επιλυθούν τέτοιου είδους προβλήματα, ακολουθούνται δύο διαφορετικές μεθοδολογίες σε ένα ταξινομητή διανυσμάτων υποστήριξης. Η πρώτη εισάγει έναν N αριθμό "χαλαρών" μεταβλητών στην αρχική εξίσωση, η οποία μετασχηματίζεται όπως φαίνεται παρακάτω:

$$y_i(w * x_i + b) \geq 1 - \xi_i, \quad i = 1, \dots, N \quad (3.5.15)$$

Ο νέος αυτός ταξινομητής διανυσμάτων υποστήριξης με την εισαγωγή "χαλαρών" μεταβλητών επιτρέπει σε ένα σημείο να βρίσκεται κατά μια μικρή απόσταση ξ στη λάθος πλευρά του υπέρ-επίπεδου χωρίς να παραβιάζεται ο περιορισμός. Συνεπώς με την χαλάρωση των περιορισμών σχετικά με το μέγιστο όριο, επιτρέπονται λανθασμένες ταξινομήσεις μέχρι ένα βαθμό. Εάν για παράδειγμα μία τιμή x_i έχει ταξινομηθεί με ένα χαλαρό όριο $0 < \xi_i < 1$ όπως φαίνεται και στην παρακάτω εικόνα είναι σωστά ταξινομημένη. Αντίθετα, στην περίπτωση που $\xi_j \geq 1$ τότε η x_j έχει ταξινομηθεί εσφαλμένα.



Εικόνα 3.14: Αδιάσπαστο γραμμικά σύνολο δεδομένων με απεικόνιση των "χαλαρών" μεταβλητών ξ_i και ξ_j , των οποίων το μέγεθος καθορίζεται από την αντίστοιχη απόστασή τους από τα όρια της κάθε κλάσης.

Εισάγοντας τη μεταβλητή του "χαλαρού" ορίου στο παραπάνω τετραγωνικό πρόβλημα βελτιστοποίησης, οι νέες εξισώσεις περιγράφονται ως:

$$\text{ελαχιστοποίηση: } Q(w, b, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (3.5.16)$$

$$\text{περιορισμοί: } y_i(w * x_i + b) \geq 1 - \xi_i, \quad i = 1, \dots, N \quad (3.5.17)$$

Η τελευταία εξίσωση περιλαμβάνει μια παράμετρο κόστους C η οποία επιλέγεται από τον χρήστη και ελέγχει την ισορροπία μεταξύ των ορίων χαλάρωσης και του μέγιστου επιτρεπτού ορίου. Η διατύπωση ξανά της εξίσωσης με τη χρήση των πολλαπλασιαστών Lagrange $\alpha_i \geq 0$ και $\beta_i \geq 0 \forall i$:

$$Q(w, b, \xi, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i [y_i(w * x_i + b) - 1 + \xi_i] - \sum_{i=1}^N \beta_i \xi_i \quad (3.5.18)$$

και όμοια με προηγούμενως οι παράγωγοι πρέπει να είναι ίσοι με το μηδέν άρα:

$$\frac{\partial Q}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i \quad (3.5.19)$$

$$\frac{\partial Q}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (3.5.20)$$

$$\frac{\partial Q}{\partial \xi_i} = 0 \Rightarrow C = \alpha_i \beta_i = 0 \Rightarrow \alpha \leq C \quad (3.5.21)$$

που διατυπώνουν ένα διπλό πρόβλημα Lagrange:

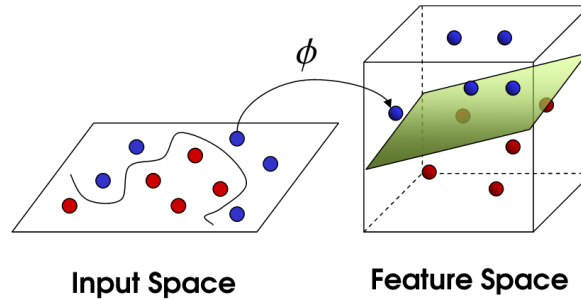
$$\text{μεγιστοποίηση: } Q(a) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j$$

$$\text{περιορισμοί: } 0 \leq \alpha_i \leq C, \forall i \quad \text{και} \quad \sum_{i=1}^N \alpha_i y_i = 0$$

Ξανά, τα διανύσματα υποστήριξης συμβολίζονται x_s , $s \in S$, όπου το S συμβολίζει το σύνολο των διανυσμάτων υποστήριξης, και επιλέγονται από τις τιμές που ικανοποιούν τα παραπάνω κριτήρια. Η συνάρτηση απόφασης "χαλαρών" ορίων ορίζεται πάλι όμοια με τη συνάρτηση απόφασης αυστηρών ορίων ως:

$$D(x) = b + \sum_{i \in S} a_i y_i x_i * x$$

Η δεύτερη μέθοδος που χρησιμοποιείται για την επίλυση ενός μη γραμμικού προβλήματος είναι η μέθοδος πυρήνα (kernel). Όπως φαίνεται και στην Εικόνα 3.15, στο μη γραμμικό πρόβλημα διαχωρισμού, δεν μπορεί να βρεθεί μία μονοδιάστατη συνάρτηση απόφασης που να είναι ικανή να διαχωρίσει τα δεδομένα σε δύο κλάσεις. Η λύση σε αυτό το πρόβλημα, όπως απεικονίζεται και στην εικόνα, έχει τη μορφή μίας κυκλικής συνάρτησης απόφασης, η οποία μπορεί να ληφθεί μονάχα σε ένα χώρο υψηλότερο διαστάσεων, όπου το σύνολο δεδομένων μπορεί να γίνει διαχωρίσιμο. Πιο συνοπτικά, η μέθοδος του πυρήνα ανάγει το πρόβλημα του διαχωρισμού, όταν αυτό είναι μη γραμμικό, μεταφέροντας τα δεδομένα σε υψηλότερη διάσταση προκειμένου να ταξινομηθούν.



Εικόνα 3.15: Μη γραμμικά διαχωρίσιμα δεδομένα στο δισδιάστατο χώρο εισόδου (αριστερά) και η προβολή τους στον τρισδιάστατο χώρο χαρακτηριστικών με χρήση συνάρτησης πυρήνα (δεξιά).

Ο χώρος των υψηλότερων διαστάσεων που μεταφέρονται τα δεδομένα ονομάζεται χώρος χαρακτηριστικών, ενώ η μεταφορά τους γίνεται με κατάλληλες συναρτήσεις. Η απευθείας αναγωγή δεδομένων σε χώρο υψηλότερων διαστάσεων με χρήση γραμμικών αλγορίθμων μάθησης, αποτελεί μια διαδικασία υψηλού υπολογιστικού κόστους, το οποίο αυξάνεται εκθετικά με την αύξηση των διαστάσεων του χώρου, γεγονός που την καθιστά μη εφικτή. Αντίθετα, η προβολή δεδομένων χρησιμοποιώντας μη γραμμικές συναρτήσεις πυρήνα, επιτρέπει την έκφραση των σημείων στον υψηλότερο διαστάσεων χώρο χαρακτηριστικών ως εσωτερικά γινόμενα. Χρησιμοποιώντας μία συνάρτηση πυρήνα $K(x_i, x_j) = \phi(x_i) * \phi(x_j)$, η συνάρτηση απόφασης διατυπώνεται ως:

$$D(x) = b + \sum_{i \in S} a_i y_i K(x_i, x_j) \quad (3.5.22)$$

Στις δημοφιλείς συναρτήσεις πυρήνα περιλαμβάνονται οι πολυωνυμικές και οι ακτινικές συναρτήσεις, όπως η συνάρτηση Gauss. Οι πρώτες διατυπώνονται ως:

$$K(x_i, x_j) = (x_i * x_j + a)^b \quad (3.5.23)$$

,ενώ αντίστοιχα μια γκαουσιανή συνάρτηση πυρήνα ορίζεται ως:

$$K(x_i, x_j) = e^{-\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)} \quad (3.5.24)$$

,όπου η τιμή σ εκφράζει την περιοχή επιρροής ενός μόνο διανύσματος υποστήριξης στη συνάρτηση λήψης απόφασης. Χαμηλές τιμές μίας μεταβλητής σ περιγράφουν μια μικρή γειτονιά ενώ υψηλές τιμές μεγαλύτερης ακτίνας περιοχή. Συνήθως, στις πειραματικές εφαρμογές ο όρος αυτός ορίζεται με τον αντίστροφο του, δηλαδή $\gamma = 1/\sigma$. Η εφαρμογή μεθόδων πυρήνα για την επίλυση τέτοιων προβλημάτων, εισάγει πρόσθετες μεταβλητές που επιλέγει ο χρήστης, όπως είναι το είδος της συνάρτησης πυρήνα ή παραμέτρους της ίδιας της συνάρτησης, όπως το πλάτος του πυρήνα σ στην περίπτωση της γκαουσιανής συνάρτησης πυρήνα. Οι μηχανές διανυσμάτων υποστήριξης διαθέτουν πολλές ιδιότητες που τις καθιστούν ιδιαίτερα δημοφιλείς ανάμεσα στους υπόλοιπους ταξινομητές.

Ένας ταξινομητής διανύσματος υποστήριξης έχει την ικανότητα να χειρίζεται μη γραμμικά διαχωρίσιμα προβλήματα, υψηλών διαστάσεων σύνολα δεδομένων [18], παρουσιάζει καλή απόδοση γενίκευσης μέσω της ελαχιστοποίησης τόσο των εμπειρικών σφαλμάτων, όσο και της πολυπλοκότητας του ταξινομητή [19], ενώ απαιτεί λίγα δεδομένα εκπαίδευσης τα οποία πρέπει να είναι αντιπροσωπευτικά [16]. Επιπλέον σε ένα ταξινομητή διανυσμάτων, απαιτούνται μικρές αλλαγές από το χρήστη μονάχα σε έναν περιορισμένο αριθμό παραμέτρων, σε αντίθεση με τα τεχνικά νευρωνικά δίκτυα, ενώ παράλληλα είναι ανθεκτικός και στην υπερφόρτωση (overfitting). Προκειμένου ο ταξινομητής διανυσμάτων να δημιουργήσει μια ισχυρή συνάρτηση απόφασης, το δείγμα εκπαίδευσης πρέπει να περιέχει έναν επαρκή αριθμό αντιπροσωπευτικών δεδομένων, τα οποία θα χρησιμοποιηθούν για την παραγωγή διανυσμάτων υποστήριξης, ωστόσο, η εκπαίδευση συνόλων δεδομένων μεγάλου μεγέθους μπορεί να αποβεί εξαιρετικά αργή.

Οι παραπάνω λόγοι κατέστησαν την χρήση των μηχανών διανυσμάτων υποστήριξης δημοφιλείς σε πολλές εφαρμογές υπολογιστικής μάθησης, σε πολλά διαφορετικά πεδία όπως είναι η αναγνώριση εικόνων και χαρακτήρων, η κατηγοριοποίηση κειμένων [20] ή ο εντοπισμός πρωτεϊνών [21]. Λαμβάνοντας υπόψη τις παραπάνω ιδιότητες επιλέχθηκαν οι μηχανές διανυσμάτων υποστήριξης ως κύριοι ταξινομητές στην παρούσα διπλωματική εργασία.

3.5.2 Μέθοδος AdaBoost

Ο όρος AdaBoost, αποτελεί ακρωνύμιο της έκφρασης ADAPtive BOOSTing , ενώ δημιουργήθηκε από τους Schaphire και Freund. Ο ταξινομητής AdaBoost ανήκει στην κατηγορία των συνδυασμένων τεχνικών υπολογιστικής μάθησης, καθώς συνδυάζει διάφορες τεχνικές υπολογιστικής μάθησης σε ένα μοντέλο πρόβλεψης με σκοπό την μείωση της διακύμανσης, της μεροληψίας και την βελτίωση των προβλέψεων. Η συγκεκριμένη αλγοριθμική διαδικασία έγινε διάσημη με την εφαρμογή του αλγορίθμου στην ανίχνευση προσώπου σε πραγματικό χρόνο που προτάθηκε τους Paula Viola και Michael Jones [22]. Η βασική αρχή του συγκεκριμένου αλγορίθμου στηρίζεται στα εξής: ένας αριθμός των λεγόμενων "αδύναμων " ταξινομητών, όπως οι απλές γραμμικές μέθοδοι απόφασης, χρησιμοποιούνται με διαδοχικό τρόπο και στη συνέχεια συνδυάζονται για τη δημιουργία ενός ταξινομητή υψηλής απόδοσης. Στο ξεκίνημα της διαδικασίας, ένας αδύναμος ταξινομητής επιχειρεί να διαχωρίσει τα δεδομένα σε κατηγορίες και να υπολογίσει τα όρια των κατηγοριών όσο το δυνατόν σαφέστερα. Εάν υποθέσουμε ότι ο όρος x_i υποδηλώνει μια τιμή από το σύνολο των δεδομένων και x_{if} ένα μοναδικό χαρακτηριστικό διάστασης f της συγκεκριμένης τιμής , η συνάρτηση απόφασης $h_t(x_i)$ με κατώφλι $th(t)$ ορίζει έναν αδύναμο ταξινομητή ως:

$$h_t(x_i) = \text{sign}(x_{if} - th(t)) \quad (3.5.25)$$

Στην περίπτωση ταξινόμησης ενός μη γραμμικά διαχωρίσιμου συνόλου δεδομένων, η εφαρμογή της μεθόδου AdaBoost ταξινομεί αρχικά λανθασμένα τα σημεία, στα οποία στη συνέχεια προσδίδει ένα βάρος a_t , όπου το t δηλώνει τον τρέχοντα γύρο. Ανάλογα με τη θέση των λανθασμένα ταξινομημένων σημείων σε σχέση με το υπέρ-επίπεδο ταξινόμησης, ο ταξινομητής τα λαμβάνει υπόψιν και του υποδεικνύεται να δώσει περισσότερη προσοχή σε αυτές τις ακραίες τιμές κατά την υλοποίηση του επόμενου $t + 1$ γύρου, έτσι ώστε να ταξινομηθούν στον επόμενο κύκλο σωστά. Τα προηγούμενα σταθμισμένα σημεία δεδομένων ταξινομούνται στη συνέχεια ξανά όσο το δυνατόν καλύτερα, αφήνοντας πάλι έναν αριθμό σημείων ακραίων τιμών, τα οποία σταθμίζονται ξανά, δηλαδή τους αποδίδονται βάρη, κατά προτεραιότητα κατά τη διάρκεια του επόμενου γύρου ταξινόμησης. Μετά το πέρας ενός ορισμένου αριθμού γύρων T που ορίζονται από το χρήστη, οι σταθμισμένοι ταξινομητές συνδυάζονται για να σχηματίσουν μία ισχυρή, μη γραμμική, επιφάνεια απόφασης:

$$D(x_i) = \text{sign}(\sum_{t=1}^T a_t h_t(x_i)), \quad D(x_i) \in (-1, +1) \quad (3.5.26)$$

,όπου $h_t(x_i \in \{-1, +1\})$ ορίζει τον ταξινομητή που εκπαιδεύτηκε στο γύρο t . Οι ταξινομητές AdaBoost απαιτούν μικρή υπολογιστική ισχύ [23] και είναι ισχυροί απέναντι στην υπερφόρτωση δεδομένων [24]. Επίσης έχουν το πλεονέκτημα να λειτουργούν με μία μέθοδο επιλογής των εγγενών χαρακτηριστικών, λαμβάνοντας υπόψη μονάχα τις διαστάσεις των χαρακτηριστικών που σχετίζονται με τη διαδικασία ταξινόμησης [25], μία ιδιαίτερα χρήσιμη ιδιότητα όταν το μέγεθος των διανυσμάτων χαρακτηριστικών είναι υψηλό. Παρόλα αυτά είναι ιδιαίτερα ευαίσθητοι σε δεδομένα που εμφανίζουν θόρυβο [26].

3.6 Αλγόριθμοι εξαγωγής και αντιστοίχισης χαρακτηριστικών

Η διαδικασία εξαγωγής χαρακτηριστικών γνωρισμάτων αποτελεί σημαντικό κομμάτι της προεπεξεργασίας πολλών αλγοριθμικών διαδικασιών και κατά καιρούς έχουν παρουσιαστεί μελέτες που ενισχύουν το κεντρικό ρόλο που διαδραματίζουν τα χαρακτηριστικά γνωρίσματα σε απαιτητικές διαδικασίες ρομποτικής όρασης. Οι διαδικασίες εξαγωγής γνωρισμάτων κατηγοριοποιούνται σε δύο ομάδες, στις ονομαζόμενες "χαμηλού " και "υψηλού " επιπέδου. Στις διαδικασίες χαμηλού επιπέδου ανήκουν όλοι οι αλγόριθμοι που μπορούν να εξάγουν χαρακτηριστικά άμεσα από μία εικόνα χωρίς καμία πληροφόρηση σε σχέση με το σχήμα του αντικειμένου, όπως είναι οι αλγόριθμοι ανιχνευτών ακμών και καμπυλών Sobel, Canny και Harris, ενώ στη δεύτερη κατηγορία ανήκουν οι τεχνικές ανίχνευσης γνωστών γεωμετρικών σχημάτων. Οι αλγοριθμικές διαδικασίες και των δύο ομάδων παρουσιάζουν αδυναμίες κατά την επίλυση εφαρμογών του φυσικού περιβάλλοντος, όπου οι σκηνές είναι αρκετά περίπλοκες, καθώς οι συνθήκες μεταβάλλονται δυναμικά, ο φωτισμός πιθανά να μην είναι ο βέλτιστος και το ένα αντικείμενο μπορεί να επικαλύπτει το άλλο κ.ά.

Η ανάγκη εξαγωγής χαρακτηριστικών γνωρισμάτων κάτω από όλες τις συνθήκες ανεξάρτητα των παραμέτρων φωτισμού, της περιστροφής ή μετακίνησης του αντικειμένου καθώς και ο εντοπισμός κάτω από οποιαδήποτε κλίμακα, με ταυτόχρονη εξάλειψη του πιθανού θορύβου στην εικόνα, οδήγησε στη δημιουργία του αλγορίθμου Scale Invariant Feature Transform (SIFT), ο οποίος αναλύεται συνοπτικά παρακάτω. Σε πολλές εφαρμογές της μηχανικής όρασης, εκτός από την διαδικασία εξαγωγής χαρακτηριστικών, εξίσου σημαντική είναι και η αντιστοίχισή τους μεταξύ των εικόνων. Όταν το ζητούμενο είναι η ανακατασκευή της σκηνής, τα χαρακτηριστικά σημεία που εξάγονται, είναι αναγκαίο να συσχετιστούν για να προκύψει η πληροφορία θέσης των πραγματικών σημείων στο χώρο, ώστε να είναι δυνατή η ανακατασκευή της γεωμετρίας της σκηνής.

Από τους πιο δημοφιλείς αλγορίθμους αντιστοίχισης σημείων είναι ο Random Sample Consensus (RANSAC), ο οποίος όταν έχει δοθεί ένα σύνολο αντιστοιχιών, έχει την ικανότητα να υπολογίζει ένα πίνακα μεταφοράς και περιστροφής που χαρακτηρίζει την πλειοψηφία των αντιστοιχισμένων σημείων. Στην παρούσα διπλωματική, οι δύο αυτές αλγοριθμικές διαδικασίες υλοποιήθηκαν για την αποκατάσταση της γεωμετρίας της σκηνής στην μεθοδολογία εξαγωγής υποβάθρου από δύο κάμερες διαφορετικού φάσματος (υπέρυθρο και ορατό), που είχαν τοποθετηθεί σε κοντινή απόσταση μεταξύ τους.

3.6.1 Αλγόριθμος SIFT

Ο αλγόριθμος Scale Invariant Feature Transform αποτελεί έναν από τους πιο δημοφιλείς αλγορίθμους εξαγωγής χαρακτηριστικών στην τεχνητή όραση και έχει αναπτυχθεί από τον David Lowe το 1999 ενώ δημοσιεύθηκε από τον ίδιο σε πιο εκτεταμένη μορφή το 2004 [27]. Ο αλγόριθμος SIFT αποτελείται από έναν ανιχνευτή χαρακτηριστικών και από έναν περιγραφέα. Στην αρχή, ο ανιχνευτής γνωρισμάτων εντοπίζει τα σημεία εκείνα που φέρουν διακριτή πληροφορία, ανεξάρτητα από μεταβολές φωτισμού, μεταφοράς, περιστροφής και κλίμακας. Με τη χρήση μιας γκαουσιανής κατανομής, η εικόνα μετασχηματίζεται από το χρωματικό χώρο κλιμάκωσης όπως προτάθηκε από τον T. Lindeberg [28]. Ακολούθως, η συνάρτηση διαφοράς του Gauss (DoG) μαζί με μία παρεμβολή στον ίδιο χώρο, παράγει τα επιθυμητά σημεία ενδιαφέροντος. Ο περιγραφέας που προαναφέρθηκε, είναι εκείνος που αναθέτει στα εξαγόμενα σημεία που προέκυψαν από τον ανιχνευτή γνωρισμάτων, προσανατολισμό. Πιο συγκεκριμένα, ο περιγραφέας είναι ένα διάνυσμα, που το μέτρο και η διεύθυνση του, υπολογίζονται από πληροφορίες της 4*4 γειτονιάς του εκάστοτε σημείου ενδιαφέροντος. Η διεύθυνσή του τυποποιείται έτσι ώστε να λαμβάνει προσανατολισμό ανά 45ο έτσι, οι τοπικές πληροφορίες ενός αντικειμένου προσδιορίζονται με έναν και μοναδικό τρόπο. Ταυτόχρονα, η απόδοση του περιγραφέα ελέγχεται συγκρίνοντας την αρχική εικόνα με την τελική και τα διανύσματα των περιγραφέων των δύο εικόνων συγκρίνονται για τυχόν ομοιότητες.

3.6.2 Αλγόριθμος RANSAC

Ο αλγόριθμος Random Sample Consensus δημοσιεύτηκε από τους Martin A. Fischler και Robert C. Bolles το 1981 [29]. Αποτελεί μια επαναληπτική διαδικασία για την εκτίμηση των παραμέτρων ενός μαθηματικού μοντέλου από ένα σύνολο παρατηρούμενων δεδομένων με ακραίες τιμές, που μπορεί να έχουν προκύψει από θόρυβο ή από λανθασμένες μετρήσεις, με σκοπό οι ακραίες τιμές να μην επηρεάζουν τις τιμές των εκτιμήσεων. Ο αλγόριθμος αυτός χρησιμοποιείται σε ένα πλήθος μαθηματικών εφαρμογών, όπως στην ανάλυση και επεξεργασία εικόνων, στην εύρεση παραμέτρων μοντέλου υπολογιστικής όρασης και στην ταξινόμηση δεδομένων. Στην περίπτωση της παρούσας διπλωματικής, το σύνολο δεδομένων ήταν οι αντιστοιχίες των χαρακτηριστικών σημείων και ζητούμενο ο υπολογισμός ενός πίνακα μεταφοράς και περιστροφής. Ο αλγόριθμος αποτελείται ουσιαστικά τα από εξής βήματα που επαναλαμβάνονται. Στο πρώτο, επιλέγεται ένα υποσύνολο που περιέχει έναν μικρό αριθμό στοιχείων των δεδομένων, τυχαία επιλεγμένα από το σύνολο των δεδομένων και υπολογίζεται ένα μαθηματικό μοντέλο. Στη συνέχεια, ο αλγόριθμος ελέγχει ποιο από τα υπόλοιπα στοιχεία του δείγματος είναι σταθερό σε σχέση με το μοντέλο, που υπολογίστηκε στο προηγούμενο βήμα. Έπειτα σημειώνει τον αριθμό των δεδομένων που μπορεί να συμπεριλάβει στο μοντέλο, βάσει μιας προκαθορισμένης τιμής και επαναλαμβάνει τα προηγούμενα βήματα για όλα τα δεδομένα. Το μοντέλο που έχει συμπεριλάβει στη λύση του τα περισσότερα δεδομένα είναι και το επικρατέστερο, ενώ τελικά υπολογίζεται το βέλτιστο μαθηματικό μοντέλο με τη βοήθεια

των ελαχίστων τετραγώνων από τα δεδομένα του επικρατέστερου μοντέλου. Όσα στοιχεία δεν είναι σταθερά θεωρούνται ακραίες τιμές που προκλήθηκαν από θόρυβο και αγνοούνται. Η διαδικασία αυτή επαναλαμβάνεται αρκετές φορές, με σκοπό την εύρεση ενός μοντέλου με τα περισσότερα σταθερά δείγματα. Ο αριθμός των επαναλήψεων που απαιτείται για τον υπολογισμό του βέλτιστου μοντέλου υπολογίζεται από την εξίσωση:

$$N = \frac{\log(1-p)}{\log(1-(1-\varepsilon)^s)} \quad (3.6.1)$$

,όπου ε η πιθανότητα το δεδομένο να αποτελεί ακραία τιμή, το s το σύνολο των δεδομένων και p ο συνολικός αριθμός των ακραίων τιμών. Στις επόμενες παραγράφους γίνεται μια σύντομη αναφορά σε ορισμένες αλγοριθμικές μεθόδους της ψηφιακής επεξεργασίας εικόνας. Οι μέθοδοι της ψηφιακής επεξεργασίας που αναλύονται είναι το μορφολογικό "άνοιγμα" και "κλείσιμο" και η ανάλυση συνδεδεμένων στοιχείων.

3.7 Ανάλυση Συνδεδεμένων Στοιχείων

Η ανάλυση συνδεδεμένων στοιχείων (Connected Components Analysis), είναι μια αλγοριθμική τεχνική της θεωρίας γραφημάτων, η οποία βρίσκει εφαρμογή στο πεδίο της μηχανικής όρασης και αποσκοπεί στην ανίχνευση και σηματοδότηση συνδεδεμένων περιοχών δυαδικών εικόνων. Ο αλγόριθμος απαιτεί ένα δομικό στοιχείο, το οποίο ορίζει τη σχέση της χωρικής γειτονίας, βάσει της οποίας εξετάζεται η τιμή του ζητούμενου εικονοστοιχείου, προκειμένου να εξεταστεί αν υφίσταται σύνδεση. Εάν υπάρχουν γειτονικά στοιχεία τα οποία έχουν σημειωθεί, το εξεταζόμενο αποκτά την μικρότερη τιμή σήμανσης που ανιχνεύθηκε στη γειτονιά του, διαφορετικά του αποδίδεται η αμέσως επόμενη από εκείνες που έχουν ήδη εκχωρηθεί. Στη δεύτερη φάση του αλγορίθμου σαρώνονται μονάχα τα δεδομένα που έχουν ήδη σηματοδοτηθεί και τους αποδίδεται η χαμηλότερη τιμή σήμανσης.

3.8 Μορφολογικές διαδικασίες Επεξεργασίας εικόνας

Οι μορφολογικοί τελεστές είναι μέθοδοι για την επεξεργασία δυαδικών εικόνων βασισμένοι στη γεωμετρία. Οι τελεστές δέχονται μια εικόνα ως είσοδο και εφαρμόζοντας ένα τελεστή, επιστρέφουν την επεξεργασμένη εικόνα ίδιων διαστάσεων με την αρχική, με την τιμή κάθε εικονοστοιχείου της εικόνας εξόδου να βασίζεται στη σχέση του εικονοστοιχείου εισόδου και του γειτονικού του. Το δομικό στοιχείο που έχει κύριο ρόλο κατά την εφαρμογή μορφολογικών τελεστών είναι αυθαίρετου σχήματος και μεγέθους, συνήθως ένα τετράγωνο 3x3. Η διάβρωση είναι μία μορφολογική διαδικασία, όπου η βασική επίδραση του τελεστή είναι η διάβρωση των ακμών μιας περιοχής εικονοστοιχείων που βρίσκονται στο προσκήνιο, με αποτέλεσμα να μικραίνουν τα αντικείμενα αφαιρώντας τα εικονοστοιχεία στα σύνορα του αντικειμένου. Αντίθετα με τη διάβρωση, η διαστολή διευρύνει τα σύνορα των περιοχών των εικονοστοιχείων που βρίσκονται στο προσκήνιο. Το μορφολογικό άνοιγμα είναι αποτέλεσμα δύο διαδικασιών, μίας διάβρωσης και μίας διαστολής, ενώ το μορφολογικό κλείσιμο, είναι αποτέλεσμα των ίδιων διαδικασιών αντίστροφα.

3.9 Βιβλιογραφική Ανάλυση: Ανίχνευση αντικειμένων μεταβαλλόμενου σχήματος

Η οπτική ανίχνευση αερίων είναι ένας σχετικά νέος τομέας έρευνας, δεδομένου του γεγονότος ότι η πρώτη κάμερα ανίχνευσης αερίων στο υπέρυθρο φάσμα εισήχθη στην αγορά μόλις το 2009. Επίσης η ανίχνευση αντικειμένων με μεταβαλλόμενη γεωμετρία αποτελεί ένα ενεργό πεδίο έρευνας με πολλά άλυτα προβλήματα, ενώ υπάρχει μόνο ένας μικρός αριθμός δημοσιευμένων μεθοδολογιών που σχετίζονται με την αυτόματη ανίχνευση και απεικόνιση αερίων. Η ενότητα αυτή παρουσιάζει μια σύντομη επισκόπηση της υπάρχουσας κατάστασης ανάπτυξης τέτοιου είδους εφαρμογών.

3.9.1 Σύνοψη συστημάτων θερμικής απεικόνισης αερίων

Οι αναφορές έρευνας που σχετίζονται με το θέμα της απεικόνισης αερίων είναι όπως αναφέρθηκε προηγουμένως αρκετά περιορισμένες. Στο άρθρο των Bouchy, Goy και Andre που παρουσιάστηκε στο Διεθνές Συνέδριο και Έκθεση για το Υγροποιημένο Φυσικό Αέριο, περιγράφεται ένα σύστημα ανίχνευσης υγροποιημένου φυσικού αερίου στο υπέρυθρο φάσμα (LNG). Το σύστημα αυτό διακρίνει το αέριο με βάση τη θερμική του υπογραφή, καθώς οι διαρροές του φυσικού αερίου τείνουν να εμφανίζουν ασυνήθιστα χαμηλές τιμές θερμοκρασίας [30]. Η συγκεκριμένη μεθοδολογία έχει εφαρμογή σε ορισμένα πεδία, καθώς εκμεταλλεύεται το γεγονός ότι μια διαρροή φυσικού αερίου παρουσιάζει διαφορά θερμοκρασίας σε σχέση με αυτή του περιβάλλοντα χώρου, στοιχείο που διαφοροποιείται από το είδος του αερίου, το μέγεθος της διαρροής και από την πίεση του δοχείου. Η εταιρεία Bertin Technologies έχει αναπτύξει ένα σύστημα ανίχνευσης αερίων που ονομάζεται Second Sight TC System, το οποίο βασίζεται στην διαφορά δύο θερμικών εικόνων υπέρυθρου φάσματος από την εφαρμογή δύο διαφορετικών φίλτρων απορρόφησης. Το πρώτο φίλτρο που βρίσκεται στην θερμική κάμερα έχει ένα ευρύ φασματικό εύρος στην οπτική διαδρομή της κάμερας και καταγράφει υπέρυθρες εικόνες της σκηνής όπου η διάχυση του αερίου δεν είναι εμφανή. Το πρόσθετο δεύτερο φίλτρο, το οποίο χαρακτηρίζεται ως "ενεργό", είναι συντονισμένο στη χαρακτηριστική ζώνη απορρόφησης του ζητούμενου αερίου και καταγράφει στιγμιότυπα όπου το αέριο είναι ορατό. Στη συνέχεια οι δύο εικόνες επικαλύπτονται, ενώ τα εικονοστοιχεία της εικόνας στα οποία είναι ορατό το αέριο που καταγράφηκε με το ενεργό φίλτρο χρωματίζονται [31], [32]. Η συγκεκριμένη μεθοδολογία είναι ικανή να διακρίνει τη διαφορά μεταξύ αερίου και αντικειμένων που δεν είναι αέρια, απαιτεί όμως τη χρήση μιας περίπλοκης και υψηλού κόστους κάμερας που περιέχει επίσης και ένα δίσκο φασματικών φίλτρων. Ένα άλλο σύστημα ανίχνευσης αερίων που λειτουργεί με πανομοιότυπο τρόπο είναι το επανομαζώσιμο Gas Vision System από την εταιρεία Gas Optics. Ο μηχανισμός αυτός στηρίζεται ξανά στην συσχέτιση δύο διαφορετικού εύρους θερμικών εικόνων από μία κάμερα για την ανίχνευση ενός αερίου στόχου, όπου αφαιρείται η μία από την άλλη προκειμένου να προκύψει μια μάσκα με μοναδικά τα εικονοστοιχεία που καταγράφουν το αέριο, η οποία στη συνέχεια τοποθετείται στην εικόνα ορατού φάσματος της σκηνής [33]. Η εταιρεία Telops, κατασκευάζει και παρέχει ένα υπερφασματικό σύστημα απεικόνισης για την ανίχνευση του αερίου που στηρίζεται στη σύγκριση των εικονοστοιχείων της καταγεγραμμένης και της χαρακτηριστικής φασματικής υπογραφής των αερίων – στόχων [34]. Οι Broadwater, Spisz και Carr πρότειναν το 2008 μια τεχνική επεξεργασίας της εικόνας για την ανίχνευση του αερίου [35]. Οι συγγραφείς του άρθρου χρησιμοποίησαν δεδομένα από στιγμιότυπα του βίντεο στο μακρύ υπέρυθρο φάσμα, ένα μοντέλο βασισμένο στους νόμους της ακτινοβολίας που είναι ικανό να εκτιμήσει τις φασματικές υπογραφές του ζητούμενου αερίου και του υποβάθρου και ένα πιθανολογικό μοντέλο απόφασης για την κατάτμηση και το χρωματισμό του αερίου. Ορμώμενοι αυτής της εργασίας, ο Gerhart και οι συνεργάτες του, παρουσίασαν μία τεχνική ξανά χρησιμοποιώντας

στιγμιότυπα βίντεο από το μακρύ υπέρυθρο φάσμα, εφαρμόζοντας τεχνικές μείωσης διαστάσεων ώστε να προβληθούν τα δεδομένα της υπερφασματικής εικόνας σε έναν ψευδή RGB χώρο, όπου στη συνέχεια πραγματοποιήσαν τεχνικές ομαδοποίησης για να διαχωρίσουν το αέριο από το υπόβαθρο της σκηνής [36]. Μια διαφορετική προσέγγιση, σε σχέση με τις προαναφερθείσες, κατοχυρώθηκε ως ευρεσιτεχνία από το ερευνητικό τμήμα της εταιρείας Shell, όπου ξανά με τις προηγούμενες τεχνικές επεξεργασίας εικόνας, αυτή τη φορά στο ορατό φάσμα επιχειρείται ο εντοπισμός του αερίου με βάση τις εξισώσεις Euler για την περιγραφή της κίνησης [37]. Ανάλογες προσεγγίσεις που στηρίζονται στον συνδυασμό διαφορών εικόνων περιγράφονται στο άρθρο των Soldan και Kroll [38], [39].

Επιπλέον, ένα σύστημα υπολογιστικής όρασης, που χρησιμοποιεί χαρακτηριστικά χρονικής συμπεριφοράς της διάχυσης των αερίων για την εποπτεία εγκαταστάσεων επεξεργασίας πετρελαίου, με την ονομασία IntelliRed αναπτύχθηκε από την εταιρεία ExxonMobil [40]. Το σύστημα αυτό εμφανίζει αρκετές ομοιότητες με το σύστημα της εταιρείας Bertin Technologies που αναλύθηκε παραπάνω, όπου γίνεται μια σύγκριση των θερμικών εικόνων καταγραφής τόσο της ύπαρξης όσο και της απουσίας αερίου από τη σκηνή προκειμένου να εξαχθούν αντικείμενα αερίου. Στη συνέχεια ακολουθεί η εφαρμογή μιας μεθοδολογίας λήψης απόφασης πολλών βαθμίδων, η οποία χρησιμοποιεί την εμφάνιση και τα δυναμικά χαρακτηριστικά του αερίου, καθώς και ένα τεχνικό νευρωνικό δίκτυο για την βελτιστοποίηση της απόδοσης της τεχνικής ανίχνευσης του αερίου [41]. Στα πλαίσια της οπτικής μηχανικής, ο D. Manolakis δημοσίευσε μία αλγοριθμική διαδικασία κατασκευής ενός στατιστικού μοντέλου για τον υπολογισμό του υποβάθρου σύμφωνα με τις θεμελιώδεις εξισώσεις της ακτινοβολίας και την επεξεργασία σήματος, το οποίο ανιχνεύει και ταυτοποιεί το χημικό αέριο που εκλύεται στο μακρύ υπέρυθρο φάσμα [42]. Όμοια με τις προηγούμενες τεχνικές, η προσέγγιση της αυτόματης ανίχνευσης αερίων που αναλύεται σε αυτή την εργασία βασίζεται τόσο σε τεχνικές υπολογιστικής όρασης και μηχανικής μάθησης, όσο και σε τεχνικές ανάλυσης εικόνας, εξαγωγής υποβάθρου και κατασκευής μάσκας. Η παρούσα διπλωματική βασίστηκε στα πλεονεκτήματα αυτών των δύο τεχνικών χρησιμοποιώντας μία θερμική κάμερα αερίου χωρίς τη χρήση εμπορικού λογισμικού και χωρίς να είναι εξοπλισμένη με πρόσθετο ενεργό αισθητήρα στην οπτική διαδρομή της.

3.9.2 Σύνοψη συστημάτων ανίχνευσης καπνού σε βίντεο

Όπως αναφέρθηκε προηγουμένως, ο αριθμός των διαθέσιμων αναφορών καθώς και των δημοσιευμένων μεθοδολογιών και τεχνικών αυτόματης ανίχνευσης αερίων είναι πολύ περιορισμένος κατά το χρόνο σύνταξης της παρούσας, καθώς η τεχνολογία αυτή βρίσκεται σε πρώιμο στάδιο έρευνας. Από την άλλη πλευρά, ένας παρεμφερής τομέας αυτόματης ανίχνευσης καπνού, φωτιάς και ατμού σε βίντεο του ορατού φάσματος, χρησιμοποιώντας τεχνικές υπολογιστικής όρασης, έχει προσελκύσει αρκετό ενδιαφέρον, ενώ παραμένει ένας τομέας διαρκούς έρευνας, με νέες προτάσεις και βελτιώσεις των υφιστάμενων τεχνικών καθώς δημοσιεύονται έρευνες σχεδόν κάθε χρόνο. Τα περισσότερα αέρια μολονότι απεικονίζονται στο υπέρυθρο φάσμα, παρουσιάζουν παρόμοια συμπεριφορά με τον καπνό σε ότι αφορά την κίνηση και την εμφάνιση, αφού εμφανίζουν και τα δύο φαινόμενα τυρβώδους ροής, χαρακτηρίζονται από έλλειψη έντονων μετατοπίσεων ενώ ταυτόχρονα χαρακτηρίζονται από απαλή υφή, χωρίς σταθερή γεωμετρία, και σε ορισμένες περιπτώσεις εμφανίζουν μερική διαφάνεια. Συνεπακόλουθα, οι τεχνικές ανίχνευσης καπνού στα στιγμιότυπα ενός βίντεο μπορούν να προσαρμοστούν για τον αυτόματο εντοπισμό αερίων. Η ανίχνευση αερίων ουσιών σε ακολουθίες βίντεο θεωρείται ως "ανοιχτό" θέμα έρευνας, δεδομένου ότι δεν υπάρχει ακόμα κάποιο μοντέλο απόλυτα επιτυχημένης ανίχνευσης λόγω ορισμένων σημαντικών προκλήσεων. Αρχικά, καθώς το σχήμα και η γεωμετρία του αερίου είναι ευμετάβλητη, η χρήση ενός ανιχνευτή περιγράμματος, όπως αυτούς που χρησιμοποιούνται κατά την ανίχνευση

αυτοκινήτων ή πεζών, δεν θα απέδιδε κάποιο αποτέλεσμα. Επίσης, τα αέρια παρουσιάζουν υψηλή μεταβλητότητα όσων αφορά στην ταχύτητα της κίνησης, την κατεύθυνση ή την στροβιλοειδή ροή, ακόμα και σε μεγέθη σχετικά με την εμφάνιση, όπως η δομή της υφής ή η διαφάνεια. Τα μεγέθη αυτά σχετίζονται με πολλούς περιβαλλοντικούς παράγοντες όπως η ταχύτητα του ανέμου, ο τύπος και η ποσότητα καμένου υλικού, στην περίπτωση εντοπισμού καπνού, καθώς και με την πίεση, το μέγεθος διαρροής, τη συγκέντρωση και τη διαφορά θερμοκρασίας με το υπόβαθρο, στην περίπτωση εντοπισμού αερίου [43].

Σε αυτή την παράγραφο γίνεται μια σύντομη επισκόπηση του δημοσιευμένου υλικού σε θέματα εντοπισμού καπνού, ατμού και πυρκαγιάς όπως δημοσιευμένα άρθρα [44] και παρουσιάσεις σε συνέδρια [45]. Οι περισσότερες προσεγγίσεις παρουσιάζουν την ίδια γενική μεθοδολογία σχετικά με την αυτόματη ανίχνευση καπνού, η οποία μπορεί να ομαδοποιηθεί σε τέσσερα βασικά στάδια, της προεπεξεργασίας, της ανίχνευσης της κίνησης για τον προσδιορισμό των κινούμενων αντικειμένων στη σκηνή, υπολογισμός των περιγραφικών χαρακτηριστικών και τέλος τη χρήση ενός ταξινομητή προκειμένου να καθορίσει αν το κινούμενο αντικείμενο είναι καπνός ή όχι. Με αυτόν τον τρόπο το πρόβλημα της ανίχνευσης του καπνού βελτιώνεται σε κάθε στάδιο, ενώ στις περισσότερες δημοσιευμένες προσεγγίσεις γίνεται εξαγωγή πολλαπλών χαρακτηριστικών ή εντάσσονται περισσότερα στάδια ανίχνευσης με σκοπό την αύξηση της ακρίβειας της ανίχνευσης. Ένα ακόμα στοιχείο που παρατηρήθηκε σε πολλές δημοσιευμένες τεχνικές είναι η κατάτμηση της εικόνας των στιγμιότυπων του βίντεο σε τετράγωνα υποπεριοχές (block), τα οποία εν συνεχεία ταξινομούνται ξεχωριστά [46]. Οι δημοσιευμένες προσεγγίσεις για την ανίχνευση καπνού, μπορούν να ταξινομηθούν είτε λαμβάνοντας υπόψη είτε μόνο τα στατικά χαρακτηριστικά εμφάνισης, ή την στατική εμφάνιση συνδυασμένη με τα χαρακτηριστικά της κίνησης.

Η πλειοψηφία των στατικών περιγραφικών χαρακτηριστικών μπορούν να χρησιμοποιηθούν για την αυτόματη ανίχνευση αερίων, με εξαίρεση εκείνους που υπολογίζουν χαρακτηριστικά με βάση το χρώμα. Οι χωροχρονικοί περιγραφείς που χρησιμοποιούνται στην ανίχνευση καπνού έχουν μοντελοποιηθεί με τέτοιο τρόπο ώστε να κωδικοποιούν την ιδιότητα που εμφανίζει ο καπνός να κινείται προς τα πάνω [47], [48], [49], μία ιδιότητα που δεν υπάρχει στην κίνηση των αερίων η οποία δεν έχει κάποια σταθερή κατεύθυνση. Επιπλέον, ξανά σε αντίθεση με τις τεχνικές εντοπισμού αερίων, στις περισσότερες δημοσιευμένες μεθοδολογίες για την αυτόματη ανίχνευση καπνού, λαμβάνεται υπόψη το χρώμα του καπνού ως στατικό χαρακτηριστικό [50], [51], [43], [46], [52], [53]. Καθώς δεν μπορούν να εξαχθούν πληροφορίες χρώματος από ένα θερμικό βίντεο και επειδή η κατεύθυνση της κίνησης ενός αερίου εξαρτάται σε μεγάλο βαθμό από το είδος του αερίου και την κατεύθυνση της διαρροής, τα δύο αυτά χαρακτηριστικά δεν εξετάζονται για την ανάπτυξη τεχνικών ανίχνευσης αερίων. Τα χαρακτηριστικά της εμφάνισης χρησιμοποιήθηκαν από τον S. Calderara και την ομάδα του [43], οι οποίοι ανέπτυξαν ένα σύστημα που αναλύει τις κινούμενες περιοχές χρησιμοποιώντας ένα προσωρινό στατιστικό μοντέλο της ενέργειας Wavelet. Στο μοντέλο αυτό εμφανίζουν υποπεριοχές της εικόνας και πληροφορίες χρώματος για να ταξινομηθούν οι περιοχές αυτές ως περιοχές ύπαρξης καπνού ή μη με έναν μπαγερσιανό ταξινομητή (Bayesian Classifier). Μία παρεμφερής δημοσίευση ανίχνευσης καπνού [52] ξανά χρησιμοποιώντας την τεχνική των Wavelets, αυτή τη φορά προκειμένου να λαμβάνει πληροφορίες σχετικά με τη διαφορά ενέργειας του υποβάθρου και του τρέχοντος στιγμιότυπου, ενώ επιπλέον λαμβάνει υπόψη πληροφορίες του χρώματος ως χαρακτηριστικά για την εφαρμογή ενός κρυφού μοντέλου Markov και ανάλυσης των ακμών.

Μια διαφορετική εφαρμογή, ανίχνευσης ατμού αυτή τη φορά, προτάθηκε από τον Z. Wang που χρησιμοποίησε μιας πλήρης κλίμακας, υψηλής ανάλυσης αποσύνθεσης κυμάτων (Wavelets) προκειμένου να εξάγει χαρακτηριστικά ομαλότητας της υφής της εικόνας και στη συνέχεια να ταξινομήσει με τη χρήση ενός ταξινομητή AdaBoost [25]. Μία ακόμη προσέγγιση που βασίζεται στην ανάλυση κυμάτων Wavelet προτάθηκε από τους J. Gubbi, S. Marusic, και M. Palaniswami, όπου τα στατιστικά χαρακτηριστικά που εξάγονται από πολλαπλά επίπεδα μετασχηματισμού των Wavelet, εκπαιδεύουν στη συνέχεια ένα ταξινομητή διανυσμάτων

υποστήριξης (SVM) [54]. Σχετικά με τα χαρακτηριστικά που βασίζονται στην κίνηση και την εμφάνιση, όπως δημοσιεύθηκε από τους K. Avgerinakis, A. Briassouli, και I. Kompatsiaris, υπολογίζονται με την ανίχνευση όλων των κινούμενων αντικειμένων, τα οποία διαχωρίζονται με τη βοήθεια εξαγωγής περιγραφών HOGHOF, με κριτήριο την κατεύθυνση της κίνησης. Στη συνέχεια τα αντικείμενα με κίνηση προς τα επάνω ταξινομούνται με έναν ταξινομητή SVM σε συνδυασμό με τον BOW (Bag of Visual Words) και τελικά εντοπίζεται ο καπνός με την εφαρμογή ενός κατωφλίου στην ενέργεια της εικόνας (Wavelets) και στις τιμές χρώματος των εικονοστοιχείων.

Βασιζόμενοι στην προηγούμενη τεχνική, οι P. Barmproutis, K. Dimitropoulos, και N. Grammalidis, πρότειναν μια παρόμοια μέθοδο που καθορίζει τις υποψήφιες περιοχές καπνού με την ανίχνευση της κίνησης και της ανάλυσης του χρώματος, ενώ στη συνέχεια βελτιώνει ένα μοντέλο απόφασης που παρακολουθεί τη χρονική συμπεριφορά της ενέργειας της εικόνας, εξάγοντας πληροφορίες με τη χρήση ξανά HOGHOF περιγραφών, του BOW αλγορίθμου και ενός ταξινομητή SVM, και εισάγοντας ταυτόχρονα ένα εκτιμώμενο διακριτό γραμμικό μοντέλο της δυναμικής της υφής του καπνού [51]. Ο Feiniu Yuan πρότεινε μία σειρά από μεθοδολογίες για τον εντοπισμό καπνού από βίντεο, όπου εκτός από την ανάπτυξη παρόμοιας τεχνικής εντοπισμού κινούμενων αντικειμένων και υπολογισμού της κατεύθυνσης τους, πρότεινε την εξαγωγή χαρακτηριστικών της υφής υποπεριοχών της εικόνας με τελεστές τοπικών δυαδικών προτύπων σε πολλαπλές κλίμακες (LBP) και ακολούθως την ταξινόμηση τους με τη χρήση ενός τεχνικού νευρωνικού δικτύου [55].

Επιπρόσθετα, ανέπτυξε άλλη μία μέθοδο χρησιμοποιώντας μία διαδικασία διπλού υπολογισμού στατιστικών ιδιοτήτων από μία σειρά πρωτογενών χαρακτηριστικών όπως είναι οι τιμές χρώματος των εικονοστοιχείων, χαρακτηριστικά προσανατολισμού ακμών (EOH) και χαρακτηριστικών δυαδικών προτύπων (LBP), τα οποία αποτέλεσαν δεδομένα εκπαίδευσης ενός ταξινομητή AdaBoost για την ταξινόμηση των χαρακτηριστικών [56].

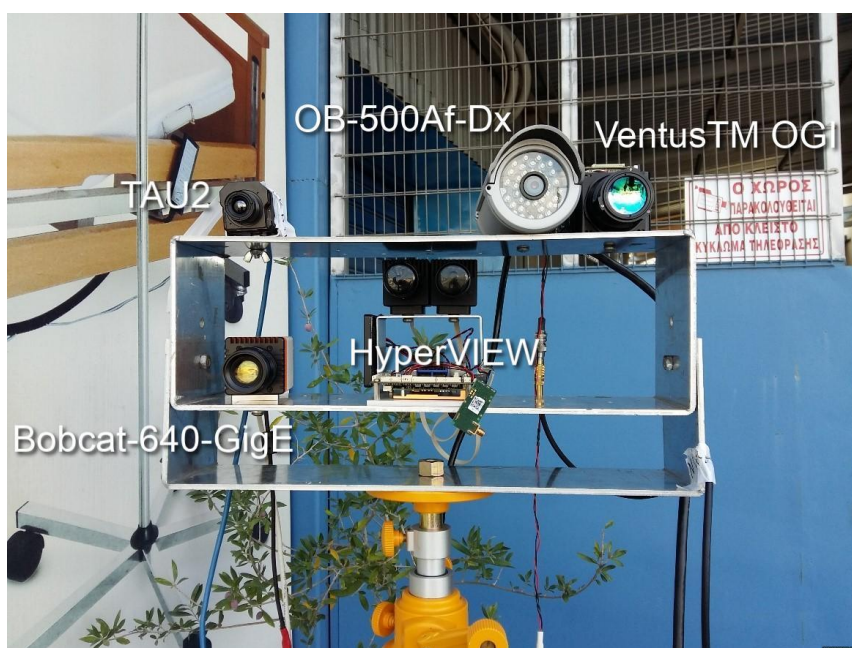
Κεφάλαιο 4^ο

Στο κεφάλαιο που ακολουθεί παρουσιάζονται αναλυτικά οι μεθοδολογίες που ακολουθήθηκαν για την αυτόματη ανίχνευση του αερίου. Αρχικά παρουσιάζεται ο εξοπλισμός με τον οποίο αποκτήθηκαν τα δεδομένα και στη συνέχεια αναλύεται η τεχνική ανίχνευσης αερίου με τη χρήση ταξινομητών. Τέλος αναλύεται η τεχνική ανίχνευσης αερίου με την εξαγωγή υποβάθρου και παρουσιάζονται τα μέτρα αξιολόγησης των δύο τεχνικών.

4 Μεθοδολογία και Δεδομένα

4.1 Εξοπλισμός και απόκτηση δεδομένων

Η απόκτηση των δεδομένων επετεύχθη με μια σύνθετη διάταξη αισθητήρων και υπολογιστικών συστημάτων. Η θερμική κάμερα που χρησιμοποιήθηκε ήταν η Ventus™ OGI της εταιρείας SIERRA OLYMPIC Technologies Inc, η οποία έχει σχεδιαστεί για την ανίχνευση και την απεικόνιση διαρροών αέριων υδρογονανθράκων όπως μεθανίου CH_4 , προπάνιου C_3H_8 , βουτανίου C_4H_{10} και άλλα. Η κάμερα Ventus™ OGI αποτελείται από έναν τελευταίας τεχνολογίας ανιχνευτή φωτονίων ευαίσθητο στις μικρές διαφορές θερμοκρασίας της σκηνής (XBn Hot Midwave IR) που λειτουργεί στο ατμοσφαιρικό παράθυρο του μεσαίου υπέρυθρου φάσματος ($3.4 - 4.2 \mu\text{m}$), ένα κλειστού κύκλου σύστημα ψύξης "stirling" και ένα συνδυασμό διαφορετικών φίλτρων απορρόφησης (ευρύ και ενεργού φίλτρου).



Εικόνα 4.1: Η πειραματική διάταξη των καμερών.



Εικόνα 4.2: Η πειραματική διάταξη (αριστερά) και τα αέρια Μεθάνιο CH_4 και προπάνιο C_3H_8 αντίστοιχα (δεξιά).

Η ανάλυση της εικόνας είναι 640×512 εικονοστοιχείων, με την απόσταση pixel pitch, δηλαδή της απόστασης μεταξύ δύο εικονοστοιχείων να είναι μόλις 1.5 χιλιοστά, ενώ ο ρυθμός καταγραφής είναι 30 στιγμιότυπα ανά δευτερόλεπτο. Καθώς οι διαστάσεις της κάμερας είναι $146.6 \times 70.9 \times 73.1$, ενώ ζυγίζει μόλις 580gr με το φακό, την καθιστά ιδανική για την ενσωμάτωση της πάνω σε Μη Επανδρωμένα Αεροσκάφη ή την τοποθέτηση της σε συστήματα συνεχούς παρακολούθησης. Η κάμερα διαθέτει αναλογικές και ψηφιακές εισόδους – εξόδους που περιλαμβάνουν "Camera Link", Gigabit Ethernet, H.264, NTSC/Pal video και RS-232/RS-422. Η ίδια εταιρεία έχει αναπτύξει ένα λογισμικό με το οποίο μπορεί να γίνει επεξεργασία των παραμέτρων του βίντεο ή επίσης εφαρμογή εργαλείων ενίσχυσης ακμών, ομαλοποίησης υφής εικόνας κ.α. Παράλληλα με την θερμική κάμερα, εγκαταστάθηκε μια απλή κάμερα που λειτουργεί στο ορατό φάσμα για την απόκτηση και σύγκριση των δεδομένων στα διαφορετικά τμήματα του ηλεκτρομαγνητικού φάσματος. Η κάμερα του ορατού φάσματος OB-500Af-Dx της εταιρείας Brickcom, κατέγραφε 24 στιγμιότυπα το δευτερόλεπτο στο ορατό φάσμα με ανάλυση εικόνας 1920×1080 εικονοστοιχεία.



Εικόνα 4.3: Παράδειγμα στιγμιότυπων raw δεδομένων (αριστερά) και δεδομένων με ενίσχυση ακμών (δεξιά) του αέριου προπάνιου.

Στην πειραματική διάταξη, που φαίνεται στην Εικόνα 4.1 υπήρχαν επιπλέον μία θερμική κάμερα, Tau2, της εταιρείας FLIR που χρησιμοποιεί ένα θερμικό ανιχνευτή με μικροβολόμετρο ευαίσθητο στο μακροκυματικό υπέρυθρο φάσμα (LWIR), μία υπερφασματική κάμερα Hyper^{VIEW} της εταιρείας 3D-ONE και της Bobcat-640-GigE της

εταιρείας Xenics, που καταγράφει βραχυκυματική υπέρυθη ακτινοβολία (SWIR). Τα δεδομένα από τις εν λόγω κάμερες δεν αξιοποιήθηκαν καθώς δεν απέβηκαν χρήσιμα για το σκοπό της παρούσας εργασίας. Όλες οι κάμερες ήταν συνδεδεμένες με εξωτερικούς φορητούς υπολογιστές, στους οποίους γινόταν η αποθήκευση των καταγραφόμενων βίντεο. Τα εξεταζόμενα αέρια στην παρούσα πειραματική διάταξη ήταν το προπάνιο και το μεθάνιο, τα οποία τα διέθεσε δωρεάν η εταιρεία πάροχος ιατρικών και βιομηχανικών αερίων REVIVAL. Η εταιρεία εκτός από τις φιάλες αερίων, παρείχε το προαύλιο του κτηρίου και τα απαραίτητα εξαρτήματα λειτουργίας των φιαλών αερίου (π.χ. ρυθμιστές πίεσης, μανόμετρα, κλπ.) για την υλοποίηση της πειραματικής διαδικασίας.



Εικόνα 4.4: Παράδειγμα στιγμιότυπων raw δεδομένων (αριστερά) και δεδομένων με ενίσχυση ακμών (δεξιά) του αερίου μεθανίου.

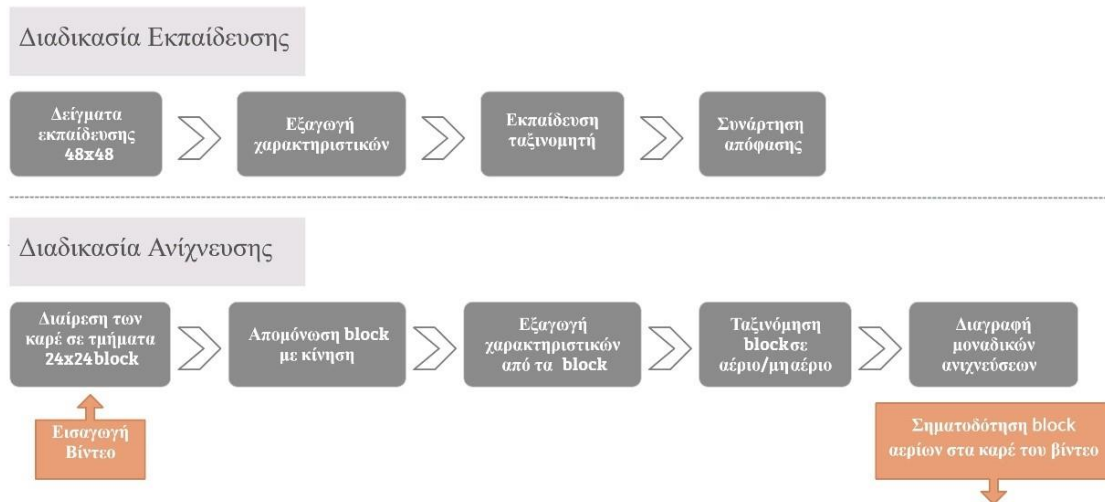
Το πείραμα υλοποιήθηκε στο πεζοδρόμιο που εκτείνεται ανάμεσα δύο τμημάτων της λεωφόρου Πειραιώς στο ύψος της περιοχής Μοσχάτου. Λόγω της θέσης υλοποίησης του πειράματος, από το οπτικό μονοπάτι της διάταξης των καμερών, διέρχονταν οχήματα, ποδήλατα και πεζοί. Επιπλέον περιλαμβανόταν ενίοτε και ο χειριστής που έλεγχε τη ροή των αερίων, καθώς δεν ήταν στάσιμος αλλά κινούνταν κατά μήκος, αντίθετα και ενδιάμεσα της ροής του αερίου. Τα παραπάνω είχαν σαν αποτέλεσμα τη δημιουργία πληθώρας θερμοκρασιακών διαφορών, σε μια δυναμική πειραματική διάταξη με μεταβαλλόμενο σκηνικό, το οποίο ανταποκρίνεται σε ένα ρεαλιστικό σενάριο διαρροής και όχι σε μια σταθερή με ελεγχόμενες συνθήκες πειραματική διαδικασία.

Επιπρόσθετα, όσον αφορά στις περιβαλλοντικές συνθήκες του εξωτερικού χώρου, επικρατούσε στη σκηνή άνεμος μέτριας έντασης και η ατμοσφαιρική θερμοκρασία είναι περίπου στους 28°C. Συνολικά δημιουργήθηκαν τέσσερα βίντεο δεδομένων, όπου στο πρώτο καταγράφεται το αέριο προπάνιο, στο δεύτερο το αέριο μεθάνιο ενώ στο τρίτο και τέταρτο γίνεται συνδυασμός των αερίων με συνεχή ροή του προπάνιου και του μεθανίου ταυτόχρονα. Τα βίντεο καταγράφηκαν με δύο τρόπους, με τον πρώτο δεν έγινε καμία διαδικασία πέρα από την καταγραφή των στιγμιότυπων από τη θερμική κάμερα, ενώ με τον δεύτερο ακολουθήθηκε μια διαδικασία ενίσχυσης ακμών μέσω του λογισμικού της εταιρείας με της κάμερας. Στην παρούσα διπλωματική υλοποιήθηκαν και οι δύο μεθοδολογίες αποκλειστικά από τα raw δεδομένα της κάμερας αλλά μόνο από το βίντεο καταγραφής της ροής του προπάνιου, καθώς το αέριο μεθάνιο μπορούσε ελάχιστα να ανιχνευθεί χωρίς την ενίσχυση ακμών από το λογισμικό της κάμερας.

4.2 Τεχνική ανίχνευσης αερίου με χρήση ταξινομητών

Όπως αναλύθηκε και στην ενότητα 3.8.1, η τεχνική ανίχνευσης αερίων με χρήση ταξινομητών είναι μια αρκετά δημοφιλής τεχνική. Στη παράγραφο αυτή αναλύεται ο τρόπος εκπαίδευσης

των ταξινομητών, η δημιουργία συνόλου δεδομένων εκπαίδευσης καθώς και η μετα-επεξεργασία των αποτελεσμάτων για την απαλοιφή ψευδών ανιχνεύσεων.



Εικόνα 4.5: Μεθοδολογία ανίχνευσης αερίου με χρήση ταξινομητών.

4.2.1 Δημιουργία δεδομένων εκπαίδευσης

Το πρώτο βήμα για την δημιουργία δεδομένων εκπαίδευσης ήταν ο διαχωρισμός των raw στιγμιότυπων βίντεο από τα επεξεργασμένα με την τεχνική ενίσχυσης ακμών. Τα raw στιγμιότυπα δεδομένα ήταν περιορισμένα σε αριθμό σε σχέση με τα επεξεργασμένα και αφορούσαν στη ξεχωριστή ροή των δύο αερίων και όχι τον συνδυασμό τους. Μετά τον διπλό διαχωρισμό των δεδομένων σε raw και επεξεργασμένα και ανάλογα με το είδος του αερίου, δημιουργήθηκαν δύο ειδών δεδομένα εκπαίδευσης. Καθώς η μεθοδολογία αφορούσε την εξαγωγή στατικών και δυναμικών χαρακτηριστικών από τους αντίστοιχους περιγραφείς, δημιουργήθηκαν δύο σύνολα θετικών και αρνητικών δεδομένων εκπαίδευσης.

Συνολικά δημιουργήθηκαν 2798 θετικά και 4687 αρνητικά στατικά δείγματα εκπαίδευσης που αφορούσαν τμήματα των στιγμιότυπων του βίντεο μεγέθους 48x48 (patches). Το μέγεθος αυτό επιλέχθηκε με γνώμονα τη δυνατότητα εξαγωγής χαρακτηριστικών από κελιά (cells), μέχρι 16x16, διαφορετικού μεγέθους για την εύρεση του βέλτιστου μεγέθους των κελιών για την κατασκευή ισχυρών περιγραφέων. Καθώς η χειροκίνητη επιλογή στα στιγμιότυπα του βίντεο θετικών και αρνητικών τμημάτων θα ήταν μια επίπονη και χρονοβόρα διαδικασία, διαιρέθηκε το σύνολο των στιγμιότυπων σε τμήματα (patches) ίδιου μεγέθους, και στη συνέχεια επιλέχθηκαν τα πιο αντιπροσωπευτικά θετικά και αντίστοιχα αρνητικά δεδομένα εκπαίδευσης. Ο συνολικός αριθμός στιγμιότυπων του raw βίντεο ανίχνευσης της ροής προπάνιου ήταν 3800 με 25 στιγμιότυπα ανά δευτερόλεπτο, ενώ κάθε στιγμιότυπο ήταν μεγέθους 352x240 εικονοστοιχείων.

Επιπρόσθετα δημιουργήθηκε ένα δεύτερο σύνολο δεδομένων, για την δημιουργία χωροχρονικών περιγραφέων χαρακτηριστικών, το οποίο αποτελούνταν από κυβοειδή δεδομένα διαστάσεων 48 εικονοστοιχείων x 48 εικονοστοιχείων x 11 στιγμιότυπα. Συνολικά κατασκευάστηκαν 102 θετικά και 160 αρνητικά κυβοειδή που απεικόνιζαν τη ροή του αερίου προπάνιου, κινούμενα οχήματα, διερχόμενους πεζούς κ.α. Η δημιουργία των κυβοειδών έγινε χειροκίνητα έπειτα από επιλογή των πιο αντιπροσωπευτικών αλληλουχιών στιγμιότυπων.



Εικόνα 4.6: Ενδεικτικά patches εκπαίδευσης με αέριο (αριστερά) και χωρίς αέριο (δεξιά) για την εξαγωγή στατικών περιγραφικών χαρακτηριστικών.

4.2.2 Προεπεξεργασία δεδομένων

Για την παραγωγή στατικών περιγραφικών χαρακτηριστικών, το πλάτος των στιγμιότυπων του βίντεο μειώθηκε στα 336 εικονοστοιχεία, με απαλοιφή εικονοστοιχείων όμοια και στις δύο κατευθύνσεις, για τη διαίρεση του κάθε στιγμιότυπου σε 140 κελιά μεγέθους 24x24 εικονοστοιχείων. Κελιά μικρότερου μεγέθους (16x16) εξετάστηκαν επίσης, αλλά ήταν πολύ μικρό το μέγεθος για την εξαγωγή χαρακτηριστικών ιδιαίτερα από του τελεστές τοπικών δυαδικών προτύπων LBP. Επιπλέον, με κελιά μεγαλύτερου μεγέθους, όπως για παράδειγμα 32x32 εικονοστοιχείων, αποτέλεσμα ήταν η εξαγωγή χαρακτηριστικών ενός μεγαλύτερου τμήματος της εικόνας χωρίς να είναι δυνατή η ανίχνευση μικρότερων περιοχών ύπαρξης του αερίου. Μετά από πολλαπλές δοκιμές τεχνικών αφαίρεσης υποβάθρου σκηνής, μέσω στατιστικών μεθόδων (Gaussian) ή μέσω επιλεκτικά μεταβαλλόμενου μοντέλου, τελικά επιλέχθηκε ο συνδυασμός των δύο πιο πάνω καθώς μοντελοποιούσε καλύτερα το υπόβαθρο της σκηνής και αφαιρούσε ικανοποιητικά σχεδόν το σύνολο των κινούμενων αντικειμένων πλην της ροής του αερίου προπάνιου. Η τεχνική του μεταβαλλόμενου μοντέλου παρουσίαζε ένα πρόβλημα, σύνηθες για τις περισσότερες τεχνικές αφαίρεσης υποβάθρου. Αυτό ήταν η εσφαλμένη συμπερίληψη αντικειμένων στο παρασκήνιο όταν έπαυαν να κινούνται για μεγάλο χρονικό διάστημα.

Στη συνέχεια εφαρμόστηκαν μια σειρά από αλγοριθμικές διαδικασίες όπως, μορφολογικό άνοιγμα, κλείσιμο και ανάλυση συνδεδεμένων στοιχείων με ταυτόχρονη απαλοιφή των σχημάτων μικρού μεγέθους, και τέλος απαλείφθηκαν τα αντικείμενα που βρέθηκαν ότι ανήκουν στο προσκήνιο αλλά εμφανίζονται σε μεμονωμένα στιγμιότυπα ή σε μία αλληλουχία πέντε ή δέκα στιγμιότυπων, λαμβάνοντας υπόψιν πως στο βίντεο αντιστοιχούσαν εικοσιπέντε στιγμιότυπα ανά δευτερόλεπτο.

4.2.3 Περιγραφείς Χαρακτηριστικών

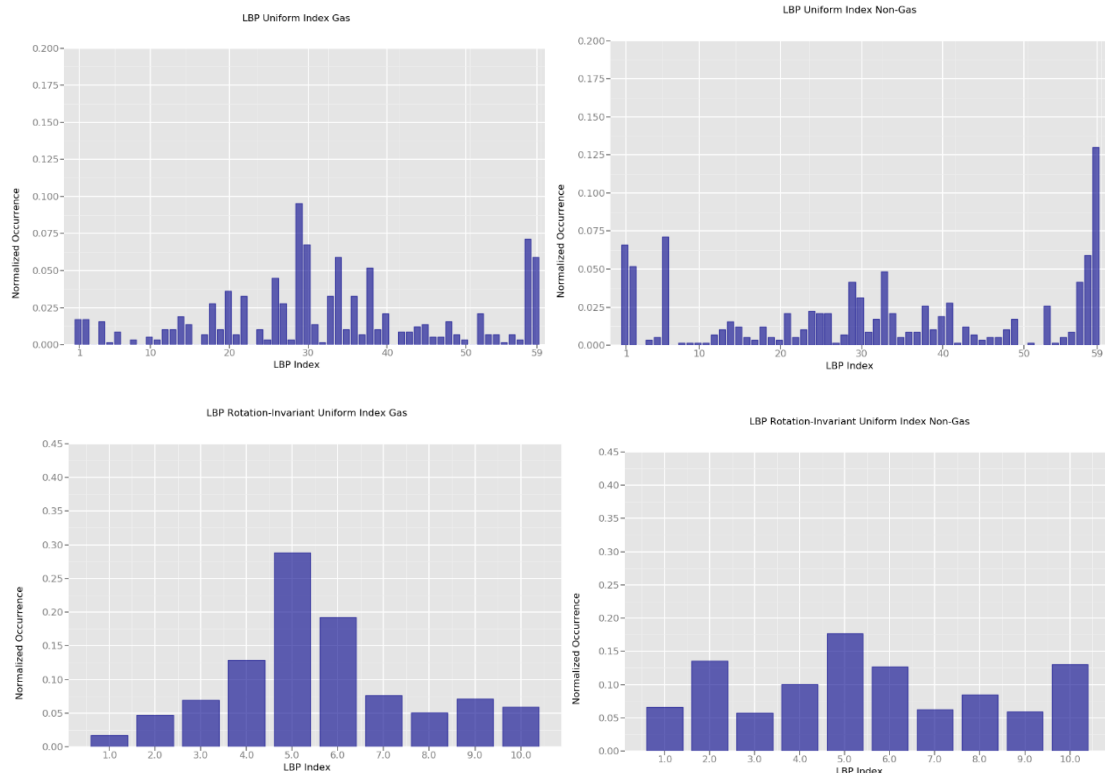
Το επόμενο στάδιο μετά την αφαίρεση του υποβάθρου ήταν η κατηγοριοποίηση των εναπομεινάντων κινούμενων αντικειμένων σε αέρια και μη αέρια, η διαδικασία που

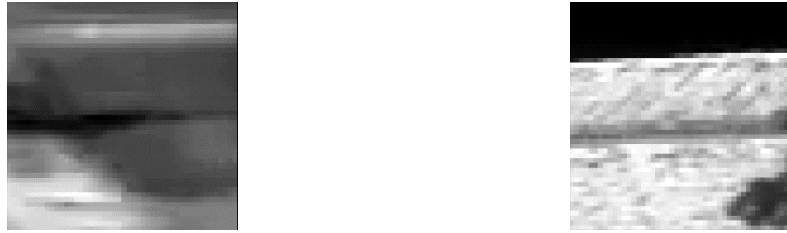
υλοποιήθηκε με την εξαγωγή περιγραφικών χαρακτηριστικών. Όπως έχει αναφερθεί, η δυσκολία εξαγωγής ισχυρών χαρακτηριστικών σχετικά με την εμφάνιση και τη συμπεριφορά των αερίων οδήγησε στην επιλογή των συγκεκριμένων περιγραφών, οι οποίοι αναλύθηκαν στο 3^ο Κεφάλαιο. Στόχος της δημιουργίας περιγραφών ήταν η αναγνώριση των χαρακτηριστικών εκείνων που ταυτοποιούν ένα αντικείμενο ως αέριο και το διαχωρίζουν από τα υπόλοιπα αντικείμενα τα οποία μπορεί να είναι παρόμοια σε εμφάνιση και κίνηση. Γενικά, τα αέρια χαρακτηρίζονται από χαμηλή αντίθεση φωτεινότητας, έλλειψη οξέων ακμών και γενικό προσανατολισμό (δηλαδή τυχαίες γεωμετρικές σχημάτων που μεταβάλλονται γρήγορα).

Η απεικόνιση των αερίων κυμαίνεται από "μόλις" ορατή εάν οι θερμοκρασιακές διαφορές στο υπόβαθρο είναι χαμηλές, έως σχεδόν αδιαφανής. Σε μικρές συγκεντρώσεις το αέριο εμφανίζεται διαφανές ή σχεδόν διαφανές, θολώνοντας έντονες γραμμές που μπορεί να υπάρχουν στη σκηνή. Αντίθετα σε υψηλές συγκεντρώσεις το αέριο εμφανίζεται σχεδόν ομοιόμορφα σε ένα σώμα ίδιας απόχρωσης. Η δυναμική της κίνησης του αερίου είναι σε μεγάλο βαθμό ασταθής και ταχέως μεταβαλλόμενη καθώς τα αέρια επηρεάζονται από την κίνηση του ανέμου.

Περιγραφέας τοπικών δυαδικών προτύπων

Προκειμένου να εξαχθούν τα χαρακτηριστικά της ομαλής ακανόνιστης υφής των αερίων κατασκευάστηκαν δύο είδη τελεστών των τοπικών δυαδικών προτύπων, ο ανεξάρτητος περιστροφής LBP^{u2} και ο ανεξάρτητος περιστροφής ομοιόμορφος τελεστής LBP^{riu2} . Οι τελεστές αυτοί είναι ιδιαίτερα ανθεκτικοί στις παραλλαγές φωτισμού, ένα χαρακτηριστικό ιδιαίτερα σημαντικό στην ανίχνευση αερίων, τα οποία αλλάζουν συνεχώς κατεύθυνση και διαφάνεια. Καθώς το μέγεθος των κελιών που επιλέχθηκε ήταν αρκετά μικρό (24x24 εικονοστοιχεία), δημιουργούνται ισχυροί περιγραφείς LBP σε γειτονιά $P=8$ και ακτίνα $R=1$.

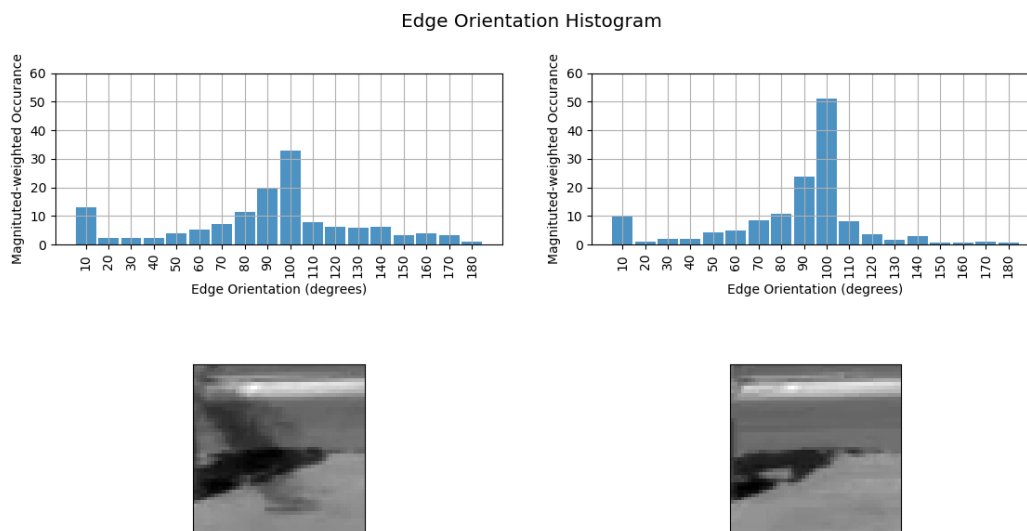




Εικόνα 4.7: Παραδείγματα ιστογραμμάτων των τελεστών LBP^{u2} (πάνω) και των LBP^{riu2} (μέση) για τα αντίστοιχα patch αερίου(κάτω αριστερά) και μη αερίου (κάτω δεξιά).

Περιγραφέας Ιστογράμματος προσανατολισμού ακμών

Οι περιοχές που καλύπτονται από τη ροή του αερίου γενικά παρουσιάζουν μικρότερα μεγέθη κλίσης, δεν υπάρχει ένας σταθερός προσανατολισμός ενώ οι κλίσεις αντικειμένων εκτός του αερίου, εμφανίζονται πιο διατεταγμένα. Για να δημιουργηθούν ισχυροί περιγραφείς χαρακτηριστικών προσανατολισμού ακμών, αξιολογήθηκαν ιστογράμματα διαφορετικών βημάτων με $K=9$, $K=18$ και $K=36$ και τελικά επιλέχθηκε η δημιουργία ιστογραμμάτων με βήμα $K=18$.

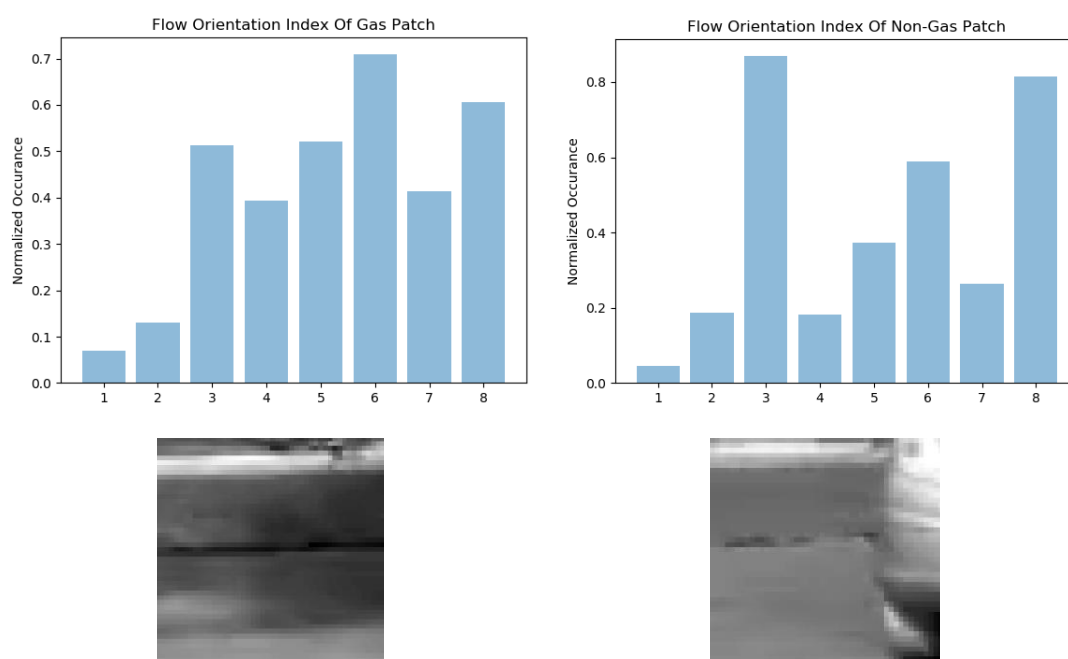


Εικόνα 4.8: Ιστογράμματα προσανατολισμού ακμών αερίου (αριστερά) και μη αερίου (δεξιά) με βήμα $K=18$ και ανά 10° .

Περιγραφέας Ιστογράμματος Οπτικής Ροής

Συγκρίνοντας τα ιστογράμματα οπτικής ροής του αερίου και ενός στερεού αντικείμενου είναι εμφανές ότι στο ιστόγραμμα του αερίου δεν υπάρχει μια γενική κατεύθυνση κίνησης που να το χαρακτηρίζει, ενώ αντίθετα το ιστόγραμμα του σταθερού αντικείμενου παρουσιάζει έντονη κορυφή που αντιστοιχεί στη γενική κατεύθυνση της κίνησης. Η ιδέα πίσω από τη χρήση του περιγραφέα χαρακτηριστικών οπτικής ροής ήταν η προσπάθεια εξαγωγής χαρακτηριστικών υφής σε συνδυασμό με την χαοτική κίνηση του αερίου. Γενικά, τα χωροχρονικά χαρακτηριστικά είναι από τη φύση τους δυσκολότερα να επεξεργαστούν, συγκριτικά με τους περιγραφείς στατικών χαρακτηριστικών, καθώς δεν αφορούν σε δυσδιάστατες εικόνες αλλά

ακολουθίες εικόνων με συνεπακόλουθη αύξηση των εξεταζόμενων διαστάσεων. Ένας τέτοιος χωροχρονικός περιγραφέας είναι η ένωση του περιγραφέα ιστογράμματος κλίσεων διαβαθμίσεων και του περιγραφέα ιστογράμματος οπτικής ροής (HOGHOF) [57]. Όπως υποδηλώνει το όνομα, οι χωροχρονικοί περιγραφείς HOGHOF εξάγουν πληροφορίες τόσο σχετικά με την εμφάνιση όσο και με τη δυναμική του τμήματος της εικόνας λαμβάνοντας υπόψη και το εξαγόμενο ιστόγραμμα κλίσεων διαβαθμίσεων αλλά και τα διανύσματα του προσανατολισμού και του μεγέθους της οπτικής ροής. Τα ιστογράμματα των δύο περιγραφέων συλλέχθηκαν από τα παραγόμενα κυβοειδή δεδομένα εκπαίδευσης, το μέγεθος των οποίων επιλέχθηκε για τις ανάγκες της παρούσας διπλωματικής, και στη συνέχεια συνενώθηκαν για να δημιουργηθούν τα HOGHOF ιστογράμματα. Η διαδικασία αυτή μπορεί να ακολουθηθεί και για διαφορετικές αναλύσεις δειγματοληψίας, όπου υψηλότερη δειγματοληψία επιτυγχάνεται όταν δεν εξάγεται ένα ιστόγραμμα περιγραφέα για κάθε ένα στιγμιότυπο του κυβοειδούς, αλλά διαιρείται το κάθε ένα στιγμιότυπο σε μικρότερες περιοχές. Υπογραμμίζεται πως η οπτική ροή δεν μπορεί να υπολογισθεί για το πρώτο στιγμιότυπο του κυβοειδούς, καθώς χρειάζονται τουλάχιστον δύο διαδοχικά στιγμιότυπα για τον υπολογισμό της ροής.



Εικόνα 4.9: Ιστογράμματα οπτικής ροής patch αερίου (αριστερά) και patch μη-αερίου (δεξιά).

Περιγραφέας Τοπικών Δυναδικών Προτύπων Τριών Διαστάσεων

Οι στατικοί περιγραφείς τοπικών δυναδικών προτύπων παρουσιάζουν καλή διακριτική ικανότητα, αλλά αποδίδουν ένα μεγάλο αριθμό ψευδών ανιχνεύσεων για αντικείμενα που έχουν παρόμοια λεία υφή. Στην παρούσα μεθοδολογία γίνεται επέκταση του τελεστή δυναδικών προτύπων από τις δύο στις τρεις διαστάσεις, που ονομάζεται περιγραφέας τοπικών δυναδικών προτύπων σε τρία ορθογώνια πλαίσια (Local Binary Patterns on Three Orthogonal Planes) και επεκτείνει την τεχνική των LBP στη χρονική διάσταση, περιγράφοντας τη δυναμική υφής όπως εξετάστηκε και προηγουμένων με τον περιγραφέα HOGHOF. Η δυναμική της κίνησης αποτελεί ένα από τα πιο ενδιαφέροντα χαρακτηριστικά περιγραφής της ροής των αερίων αλλά παρουσιάζει χαμηλή διακριτική ικανότητα αν δεν συνδυαστεί κατάλληλα με έναν περιγραφέα χαρακτηριστικών υφής. Ο περιγραφέας LBPTOP υλοποιήθηκε στην παρούσα τεχνική καθώς αποτελούσε μια προέκταση του ισχυρού τελεστή χαρακτηριστικών LBP, ενώ για την

κατασκευή του χρησιμοποιήθηκε ο ανεξάρτητος περιστροφής LBP^{u2}, με μέγεθος ιστογράμματος 59, στα τρία επίπεδα δημιουργώντας έναν διάνυσμα χαρακτηριστικών μήκους 177. Οι τελεστές LBP αξιολογήθηκαν για κυβοειδή διαφορετικού βάθους διαστάσεων 24 εικονοστοιχείων x 24 εικονοστοιχείων x W_i με $i \in \{7,9,11\}$ αριθμός διαδοχικών στιγμιότυπων. Ιδιαίτερη προσοχή δόθηκε στην επιλογή του αριθμού των διαδοχικών στιγμιότυπων προς εξέταση καθώς σε επιλογή μεγαλύτερου αριθμού ενδέχεται ο κίνδυνος εξαγωγής χαρακτηριστικών από στιγμιότυπα που δεν εμφανίζεται το αέριο.

4.2.4 Ταξινόμηση των περιοχών αερίου

Αφότου παρήχθησαν όλοι οι διαφορετικοί περιγραφείς χαρακτηριστικών, στατικοί και δυναμικοί, χρησιμοποιήθηκε το σύνολο των θετικών και αρνητικών δειγμάτων ως δεδομένα εισόδου εκπαίδευσης ταξινομητών. Στην παρούσα έρευνα πραγματοποιήθηκε εφαρμογή δύο διαφορετικών ταξινομητών για τη σύγκριση της απόδοσης της διαχωριστικής τους ικανότητας, ενός ταξινομητή μηχανών διανυσμάτων υποστήριξης (SVM) και ενός ταξινομητή AdaBoost. Όπως αναφέρθηκε και στην ενότητα 3.5.1 με χρήση των ταξινομητών SVM επιτυγχάνεται υψηλή διακριτική ικανότητα όταν τα δείγματα είναι αντιπροσωπευτικά και ικανοποιητική απόδοση ακόμα και σε περίπτωση μικρού αριθμού δεδομένων εκπαίδευσης. Η διαδικασία που ακολουθήθηκε για την εκπαίδευση του ταξινομητή ήταν η ακόλουθη:

- i. Φόρτωση των θετικών και των αρνητικών δειγμάτων patches εικόνων.
- ii. Υπολογισμός των περιγραφών χαρακτηριστικών και προσθήκη θετικής (+1) ή αρνητικής (-1) τιμής ως ετικέτα.
- iii. Ανάμειξη όλων των διανυσμάτων χαρακτηριστικών με τις αντίστοιχες ετικέτες τους και διαχωρισμός των δεδομένων σε ένα σετ εκπαίδευσης και ένα σετ δοκιμών
- iv. Δοκιμή του εκπαιδευμένου ταξινομητή με το σετ δοκιμών για την αξιολόγηση της διακριτικής του ικανότητας σε νέα άγνωστα δεδομένα.

Η δοκιμή του ταξινομητή μόνο με τα εξαγόμενα patches του σετ δοκιμών, τα οποία είχαν δημιουργηθεί ως δεδομένα εκπαίδευσης, αποτελεί επιεική εκτίμηση της απόδοσης του ταξινομητή. Για μια αυστηρότερη αξιολόγηση όλης της μεθοδολογίας ανίχνευσης με τη χρήση του ταξινομητή για τη διάκριση των περιοχών της εικόνας σε αέριο και μη-αέριο, ο εκπαιδευμένος ταξινομητής να εξετασθεί στις ακολουθίες των στιγμιότυπων του βίντεο. Επιπλέον είναι σημαντικό το σετ εκπαίδευσης να μην επικαλύπτεται με το σετ δοκιμών, δηλαδή τα δείγματα που συντέλεσαν στην εκπαίδευση του ταξινομητή να μην χρησιμοποιούνται στο σετ δοκιμών για την αξιολόγηση της απόδοσης. Καθώς όλα τα δεδομένα εκπαίδευσης στη παρούσα εργασία προέρχονται από ένα μοναδικό βίντεο ανίχνευσης του αερίου προπάνιου μία πιο αυστηρή αξιολόγηση της απόδοσης του ταξινομητή δεν ήταν εφικτή λόγω του μικρού αριθμού διαφορετικών δειγμάτων εκπαίδευσης. Για κάθε διαφορετικό περιγραφέα χαρακτηριστικών δημιουργήθηκε ένας ταξινομητής και στη συνέχεια αποθηκεύτηκε σε ένα αρχείο για μεταγενέστερη επεξεργασία του.

Ο ταξινομητής AdaBoost χρησιμοποιήθηκε για τη δημιουργία συμπαγών διανυσμάτων χαρακτηριστικών χρησιμοποιώντας τη μέθοδο της εγγενούς επιλογής χαρακτηριστικών. Για τη σύγκριση της απόδοσης των δύο ταξινομητών χρησιμοποιήθηκαν οι ίδιοι περιγραφείς που χρησιμοποιήθηκαν και κατά την εκπαίδευση του ταξινομητή SVM, προκειμένου να επιλεχθεί εκείνος που μεγιστοποιεί την απόδοση της ανίχνευσης. Ο αριθμός των κύκλων ενίσχυσης τέθηκε σε $T=100$ καθώς παρατηρήθηκε ότι το σφάλμα εκπαίδευσης παρέμενε σταθερό για μεγαλύτερο αριθμό. Μετά την ολοκλήρωση της δημιουργίας και των δύο διαφορετικών ειδών ταξινομητών, ακολούθησε η δημιουργία των βίντεο πρόβλεψης. Από την εξαγωγή του υποβάθρου που αναλύθηκε στην ενότητα 4.2.2 βρέθηκαν οι περιοχές ενδιαφέροντος για την ταξινόμηση τους σε περιοχές ύπαρξης αερίου ή μη. Αφού διαιρέθηκαν τα στιγμιότυπα του θερμικού βίντεο δεδομένων σε 140 κελιά, ερευνήθηκε από το βίντεο υποβάθρου ποια ήταν τα

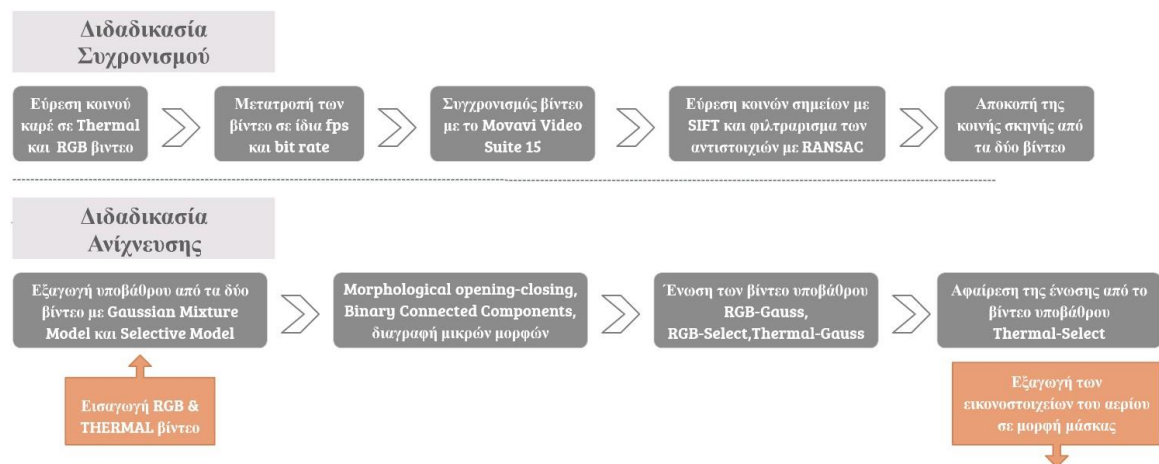
κελιά τα οποία περιείχαν κίνηση αντικειμένων προκειμένου να εξαχθούν οι περιγραφείς χαρακτηριστικών και στη συνέχεια να ταξινομηθούν σε κελιά ύπαρξης αερίου ή μη.

4.2.5 Μετά-Επεξεργασία Αποτελεσμάτων

Προκειμένου να μειωθεί ο αριθμός των ψευδών ανιχνεύσεων εκτιμήθηκε η χρονική εξέλιξη των θετικών προβλέψεων ανίχνευσης στα στιγμιότυπα του βίντεο. Πιο αναλυτικά, από το βίντεο των προβλέψεων, το οποίο περιείχε κελιά θετικών προβλέψεων, μελετήθηκε η εξέλιξη ενός κελιού πρόβλεψης στα επόμενα στιγμιότυπα του βίντεο πρόβλεψης, όπου μεμονωμένες ανιχνεύσεις σε ένα στιγμιότυπο ή δύο διαδοχικά στιγμιότυπα θεωρήθηκαν εσφαλμένες, καθώς η ροή του αερίου θεωρείται ότι να είναι συνεχόμενη. Τέλος, τα εναπομείναντα κελιά σημειώθηκαν στο τελικό βίντεο αποτελεσμάτων με χρώμα προκειμένου να γίνεται αντιληπτή η ανίχνευση του αερίου.

4.3 Τεχνική ανίχνευσης αερίου με αφαίρεση υποβάθρου

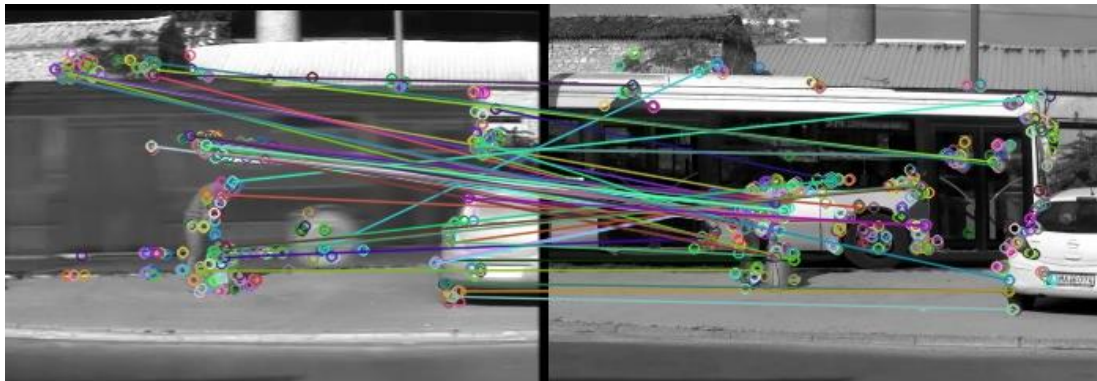
Η τεχνική ανίχνευσης με συνδυασμό στιγμιότυπων από βίντεο διαφορετικών περιοχών ηλεκτρομαγνητικού φάσματος αποτελεί μια πρωτότυπη τεχνική. Η μέθοδος αυτή στηρίζεται στο γεγονός ότι η καταγραφή της ίδιας σκηνής σε διαφορετικά μήκη του φάσματος δίνει τη δυνατότητα αφαίρεσης του κοινού υποβάθρου και απομόνωσης μόνο της περιοχής ενδιαφέροντος που δεν είναι ορατή σε ένα φάσμα. Στην ενότητα αυτή αναλύεται η μεθοδολογία που ακολουθήθηκε η οποία περιλάμβανε τον συγχρονισμό των δύο βίντεο, την απομόνωση της κοινής σκηνής στις καταγραφόμενες εικόνες, τον συνδυασμό των τεχνικών αφαίρεσης υποβάθρου που χρησιμοποιήθηκαν καθώς και την μετά-επεξεργασία των αποτελεσμάτων για την απαλοιφή των ψευδών ανιχνεύσεων.



Εικόνα 4.10: Μεθοδολογία ανίχνευσης αερίου με συνδυασμό μεθόδων αφαίρεσης υποβάθρου.

4.3.1 Συγχρονισμός των βίντεο

Όπως προαναφέρθηκε, από την πειραματική διαδικασία δημιουργήθηκαν δύο βίντεο σε διαφορετικά τμήματα του ηλεκτρομαγνητικού φάσματος, ένα στο ορατό και ένα στο μεσαίο υπέρυθρο τμήμα. Τα δύο βίντεο, εκτός από την καταγραφή της ίδιας σκηνής σε διαφορετικά μήκη κύματος, διέφεραν στις διαστάσεις της εικόνας (διαφορετικό ύψος και πλάτος), στα στιγμιότυπα ανά δευτερόλεπτο (fps) καθώς και διαφορετικούς ρυθμούς μετάδοσης bit (bit rate). Ο αριθμός των στιγμιότυπων ανά δευτερόλεπτο και ο ρυθμός μετάδοσης διορθώθηκαν έτσι ώστε να ίδιοι και στα δύο βίντεο. Επιπλέον, δημιουργήθηκε διαφορετικός αριθμός θερμικών και ορατού φάσματος βίντεο, για την ίδια πειραματική διαδικασία, καθώς η κάμερα ορατού φάσματος επέτρεπε καταγραφή περιορισμένου χρόνου.



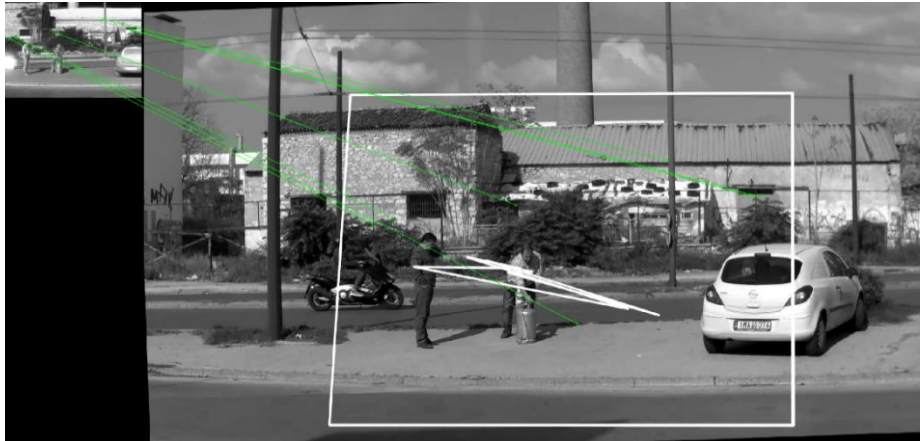
Εικόνα 4.11: Εύρεση ομόλογων σημείων με τον αλγόριθμο SIFT στις δύο εικόνες.

Συνεπακόλουθα, δημιουργήθηκαν τέσσερα RGB βίντεο που αντιστοιχούσαν στο θερμικό βίντεο καταγραφής προπάνιου, ενώ επίσης τρία RGB αντιστοιχούσαν στο βίντεο ανίχνευσης μεθανίου. Λόγω της ύπαρξης χρονικών κενών, δηλαδή έλλειψης ορισμένων στιγμιότυπων καταγραφής της σκηνής στα RGB βίντεο, έγινε αναγκαία η εύρεση κοινών στιγμιότυπων για την υλοποίηση του συγχρονισμού τους. Πιο αναλυτικά, βρέθηκε ένα κοινό στιγμιότυπο έναρξης και τέλους στα δύο βίντεο, με στόχο να απεικονίζουν στην ίδια χρονική στιγμή την ίδια σκηνή στο ξεκίνημα και στο τελείωμα του κάθε βίντεο. Ο συγχρονισμός των βίντεο υλοποιήθηκε με τη βοήθεια του εμπορικού λογισμικού Monavi Video Suite 15. Μετά την ολοκλήρωση του συγχρονισμού ακολούθησε η αποκατάσταση της γεωμετρίας της σκηνής.



Εικόνα 4.12: Η επικαλυπτόμενη περιοχή των δύο εικόνων.

Καθώς οι δύο κάμερες ήταν τοποθετημένες σε κοντινή απόσταση μόλις λίγων εκατοστών η μία από την άλλη, η καταγραφόμενη σκηνή στα δύο βίντεο διέφερε κατά μια μικρή γωνία. Επιπλέον, όπως αναφέρθηκε στις διαφορές των παραγόμενων βίντεο, τα δύο βίντεο διέφεραν στις διαστάσεις της εικόνας με αποτέλεσμα, η θερμική εικόνα να αποτελεί κομμάτι της εικόνας RGB. Προκειμένου να γίνει ακριβής αναγωγή της εικόνας με υπολογισμό της περιοχής επικάλυψης των δύο παραγόμενων εικόνων, έγινε εφαρμογή αλγοριθμικών διαδικασιών εξαγωγής και αντιστοίχισης χαρακτηριστικών. Τα ομόλογα σημεία στις δύο εικόνες βρέθηκαν με τον αλγόριθμο Scale Invariant Feature Transform (SIFT).



Εικόνα 4.13: Φιλτράρισμα αντιστοιχιών με τον αλγόριθμο RANSAC και αφινικός μετασχηματισμός της εικόνας.

Ο αλγόριθμος υλοποιήθηκε σε διαφορετικά κοινά στιγμιότυπα μέχρι την εύρεση του βέλτιστου αριθμού ομόλογων σημείων. Στη συνέχεια, ακολούθησε η διαδικασία αντιστοίχισης των ομόλογων σημείων στις δύο εικόνες με χρήση του αλγόριθμου αντιστοίχισης Fast Library for Approximate Nearest Neighbors (FLANN), που αντιστοιχεί τα σημεία με βάση τη μέθοδο του εγγύτερου γείτονα. Ακολούθως, πραγματοποιήθηκε φιλτράρισμα των αντιστοιχιών με τη χρήση του αλγορίθμου Random Sample Consensus (RANSAC), προκειμένου να βρεθούν οι καλύτερες αντιστοιχίες των ομολογιών. Συνολικά, από τις σαράντα αντιστοιχίες, επτά κρίθηκαν ως βέλτιστες και αποτέλεσαν τη βάση για την εύρεση της επικαλυπτόμενης περιοχής των εικόνων με τη ταυτόχρονη εφαρμογή αφινικού γραμμικού μετασχηματισμού και της αντίστοιχης μετατόπισης. Τέλος, μετά την αποκοπή μόνο της επικαλυπτόμενης περιοχής των δύο εικόνων, με χρήση παρεμβολής μετατράπηκε το μέγεθος των στιγμιότυπων του RGB στις διαστάσεις των εικόνων του θερμικού βίντεο.

4.3.2 Ανίχνευση αερίου με εξαγωγή υποβάθρου

Η ανίχνευση του αερίου στην τεχνική αφαίρεση του υποβάθρου, στηρίζεται σε ένα συνδυασμό τεχνικών εξαγωγής, με ένα συνδυασμό γκαουσιανών μοντέλων (Gaussian Model Mixture) και ενός επιλεκτικά μεταβαλλόμενου μοντέλου. Η τεχνική αφαίρεσης υποβάθρου με ένα γκαουσιανό μοντέλο, επέστρεψε ικανοποιητικά αποτελέσματα για την αφαίρεση μόνο των κινούμενων αντικειμένων, εκτός του αερίου, καθώς ήταν ικανή να ανιχνεύσει μόνο τις υψηλές μεταβολές στο υπόβαθρο της σκηνής, όπως κινούμενα οχήματα, πεζούς αλλά όχι την κίνηση αερίου.



Εικόνα 4.14: Το στιγμιότυπο του θερμικού με ύπαρξη αερίου (αριστερά, με κόκκινο η περιοχή ύπαρξης αερίου) και το εξαγόμενο υπόβαθρο με τη τεχνική GMM.

Αντιθέτως, η τεχνική αφαίρεσης υποβάθρου μέσω επιλεκτικά μεταβαλλόμενου μοντέλου, λόγω της ιδιότητας της ανατροφοδότησης του μοντέλου του υποβάθρου, κατέστησε τα εικονοστοιχεία του αερίου ανιχνεύσιμα στο προσκήνιο. Όπως αναφέρθηκε και στην ενότητα 3.1.3, ο προσδιορισμός των κινούμενων αντικειμένων επιτυγχάνεται με την αφαίρεση του κάθε νέου στιγμιότυπου από το αρχικό υπόβαθρο που έχει οριστεί.

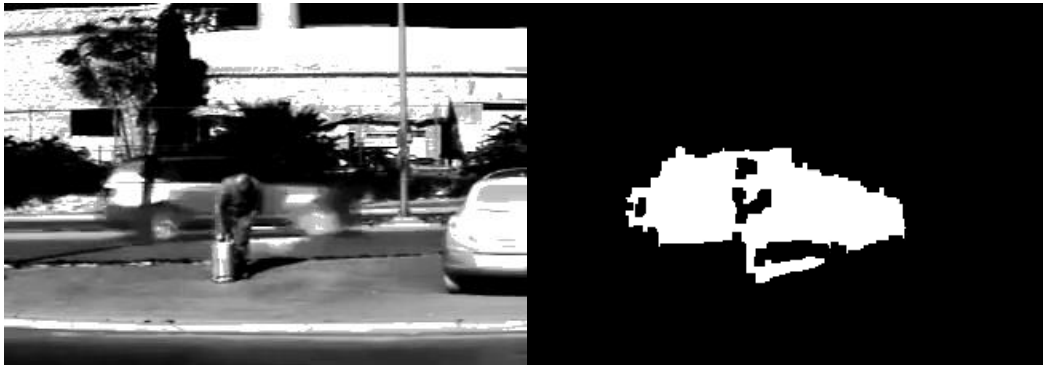


Εικόνα 4.15: Το στιγμιότυπο του θερμικού με ύπαρξη αερίου (αριστερά, με κόκκινο η περιοχή ύπαρξης αερίου) και το εξαγόμενο υπόβαθρο με τη τεχνική Selective model.

Η ανίχνευση κινούμενων αντικειμένων με αφαίρεση υποβάθρου στο ορατό τμήμα του φάσματος, απέδιδε όλα τα κινούμενα αντικείμενα πλην του αερίου. Συμπληρωματικά, η εξαγωγή υποβάθρου στο θερμικό φάσμα, εντόπιζε όλα τα ήδη εντοπισμένα κινούμενα αντικείμενα, καθώς επίσης και το αέριο. Η αφαίρεση των δύο εικόνων ανίχνευσης είχε ως αποτέλεσμα τον εντοπισμό των εικονοστοιχείων που αντιστοιχούσαν στο αέριο προπάνιο. Λόγω των διαφορετικών παραμέτρων καταγραφής της σκηνής από την κάθε κάμερα, κυρίως του διαφορετικού ρυθμού καταγραφής στιγμιότυπων, το αποτέλεσμα της μεθόδου του συγχρονισμού παρουσιάζει ορισμένες ατέλειες.

Προκειμένου να απαλειφθούν οι ατέλειες αυτές, υλοποιήθηκε ο συνδυασμός τεχνικών αφαίρεσης υποβάθρου προκειμένου να δημιουργηθεί ένα ισχυρό υπόβαθρο που είναι ικανό να περικλείσει μονάχα τα εικονοστοιχεία του αερίου. Συνολικά δημιουργήθηκαν τέσσερα διαφορετικά βίντεο εξαγωγής υποβάθρου, δύο από την εφαρμογή γκαουσσιανών μοντέλων (Gaussian Model Mixture) και δύο από τη χρήση επιλεκτικά μεταβαλλόμενων μοντέλων. Στη συνέχεια προστίθενται τα στιγμιότυπα των τριών βίντεο, στα οποία δεν πραγματοποιείται ανίχνευση των εικονοστοιχείων του αερίου στο υπόβαθρο (GMM-RGB, Selective model-RGB και GMM-Thermal), προκειμένου να δημιουργηθεί ένα βίντεο, όπου στα στιγμιότυπά του εμφανίζονται τα εικονοστοιχεία όλων των κινούμενων αντικειμένων πλην του αερίου. Πριν την ένωση των βίντεο, υλοποιείται μια σειρά από μορφολογικές διαδικασίες καθώς και

ανάλυση συνδεδεμένων στοιχείων ίδιου πυρήνα για την πλήρωση κενών των κινούμενων αντικειμένων.



Εικόνα 4.16: Το αποτέλεσμα του συνδυασμού των τριών βίντεο υποβάθρου μετά από μορφολογικές διαδικασίες (δεξιά) και το αντίστοιχο στιγμιότυπο του θερμικού βίντεο (αριστερά).

Στη συνέχεια, αφαιρέθηκαν τα στιγμιότυπα της πρόσθεσης των τριών βίντεο από τα στιγμιότυπα που προέκυψαν από την εφαρμογή του επιλεκτικά μεταβαλλόμενου μοντέλου στο θερμικό βίντεο, για την απαλοιφή των εικονοστοιχείων των υπόλοιπων κινούμενων αντικειμένων και την απομόνωση των εικονοστοιχείων του αερίου. Στο αποτέλεσμα της αφαίρεσης ξανά υλοποιείται μορφολογικό άνοιγμα πυρήνα 3×3 και στη συνέχεια κλείσιμο με πυρήνα 5×5 για την απαλοιφή μικρών εναπομεινάντων σχημάτων, ενώ έπειτα ακολουθείται ανάλυση συνδεδεμένων στοιχείων για τη σύνδεση σχημάτων που χωρίστηκαν κατά την εφαρμογή των μορφολογικών διαδικασιών.



Εικόνα 4.17: Αποτέλεσμα ανίχνευσης μετά τις μορφολογικές διαδικασίες και ανάλυσης συνδεδεμένων στοιχείων του βίντεο διαφοράς υποβάθρων (δεξιά, με κόκκινο περίγραμμα τα υπολειπόμενα σχήμα εικονοστοιχείων μη-αερίου) και το αντίστοιχο στιγμιότυπο του θερμικού βίντεο (αριστερά).

Η απαλοιφή σχημάτων που δεν αντιστοιχούσαν σε αέριο (όπως φαίνονται στην Εικόνα 4.16 με κόκκινο περίγραμμα) υλοποιήθηκε σε τρία στάδια. Αρχικά αφαιρέθηκαν σχήματα μικρού μεγέθους (200 εικονοστοιχείων) καθώς το νέφος του αερίου απεικονίζεται στο θερμικό βίντεο ως μία συμπαγής και μεγάλου μεγέθους μορφή. Στη συνέχεια και έχοντας υπόψη ότι τα σχήματα αερίου απεικονίζονται σε διαδοχικά στιγμιότυπα καθώς αφορούν σε μια συνεχή σταθερή ροή, υπολογίσθηκαν τα κέντρα των σχημάτων, για να αφαιρεθούν εκείνα που εμφανίζονταν μόνο για ένα, πέντε ή δέκα συνεχόμενα στιγμιότυπα. Υπογραμμίζεται πως το κεντροειδές ενός σχήματος, στην παρούσα τεχνική, είναι ο αριθμητικός μέσος όρος όλων των

εικονοστοιχείων που ανήκουν σε ένα σχήμα. Ο υπολογισμός των κέντρων των σχημάτων υλοποιήθηκε με βάση τα στιγμιότυπα εικόνας (Image Moments). Ορίζοντας σύντομα εδώ το στιγμιότυπο της εικόνας, αποτελεί έναν ιδιαίτερο σταθμισμένο μέσο όρο των εντάσεων των εικονοστοιχείων μιας εικόνας με τον οποίο υπολογίζονται οι ιδιότητές της όπως η ακτίνα, η περιοχή ή το κέντρο μίας μονόχρωμης εικόνας. Προκειμένου να βρεθούν τα κεντροειδή όλων των σχημάτων σε κάθε στιγμιότυπο του βίντεο εξαγωγής υποβάθρου βρέθηκαν πρώτα τα όρια (contours) κάθε σχήματος. Τέλος, μελετήθηκε η μετατόπιση των κεντροειδών στα επόμενα στιγμιότυπα, στα οποία όταν βρέθηκε ότι οι μετατοπίσεις ήταν μεγαλύτερες των είκοσι εικονοστοιχείων, θεωρήθηκε πως αντιστοιχούσαν σε κινούμενα αντικείμενα και όχι σε αέριο, όπως κινούμενα οχήματα και πεζοί. Μετά το τέλος της διαδικασίας απαλοιφής σχημάτων, δημιουργήθηκε ένα βίντεο μάσκας με τα εναπομείναντα εικονοστοιχεία που θεωρήθηκαν ότι ανήκαν στο αέριο, με τα οποία έγινε αναγωγή στο αρχικό θερμικό βίντεο με χρωματισμό των συγκεκριμένων με μία χρωματική παλέτα για την οπτικοποίηση του αερίου.

4.4 Αξιολόγηση Απόδοσης με χρήση Πίνακα Σύγχυσης

Η απόδοση ενός ταξινομητή μπορεί να υπολογισθεί με πολλούς διαφορετικούς τρόπους. Λαμβάνοντας υπόψη πως στην παρούσα εργασία εκπαιδεύτηκαν δυαδικοί ταξινομητές, οι πιθανές εκβάσεις ήταν ορθές θετικές προβλέψεις (όταν μια θετική περίπτωση είναι σωστά ταξινομημένη ως θετική), λανθασμένες θετικές προβλέψεις (όταν μία αρνητική περίπτωση είναι εσφαλμένα ταξινομημένη ως θετική), ορθές αρνητικές προβλέψεις (όταν μία αρνητική περίπτωση είναι σωστά ταξινομημένη ως αρνητική) και λανθασμένες αρνητικές προβλέψεις (όταν μία θετική περίπτωση είναι εσφαλμένα ταξινομημένη ως αρνητική). Ένας πίνακας σύγχυσης δύο προς δύο, παρέχει μια γρήγορη επισκόπηση του αριθμού των ορθών και των εσφαλμένων προβλέψεων:

Πίνακας 4-1: Πίνακας Σύγχυσης.

	Ταξινομημένα Θετικά (Classified Positive)	Ταξινομημένα Αρνητικά (Classified Negative)
Πραγματικά Θετικά (Actual Positive)	Ορθά Θετικά (True Positive)	Εσφαλμένα Αρνητικά (False Negative)
Πραγματικά Αρνητικά (Actual Negative)	Εσφαλμένα Θετικά (False Positive)	Ορθά Αρνητικά (True Negative)

Οι πιο δημοφιλείς μετρικοί παράγοντες που υπολογίζονται από έναν πίνακα σύγχυσης συνοψίζονται στα ποσοστά πληρότητας ή ανάκλησης, ορθότητας ή ακρίβειας, συνολικής ποιότητας και περιοχής κάτω από την καμπύλη ROC. Πιο αναλυτικά, το ποσοστό πληρότητας εκφράζει πόσες από το σύνολο των πραγματικών θετικών τιμών ταξινομήθηκαν σωστά, ενώ το ποσοστό της ορθότητας εκφράζει πόσες από το σύνολο των θετικών ανιχνεύσεων, αντιστοιχούσαν σε πραγματικά θετικές ανιχνεύσεις.

$$\text{Completeness (Recall)} = \frac{TP}{TP+FN} \quad (4.4.1)$$

και

$$\text{Correctness (Precision)} = \frac{TP}{TP+FP} \quad (4.4.2)$$

Δύο ακόμα δημοφιλή μέτρα που εξάγονται από έναν πίνακα σύγκρισης είναι ο ρυθμός των ορθών θετικών προβλέψεων (True Positive Rate), που εκφράζει τη συχνότητα πρόβλεψης θετικών τιμών από τον ταξινομητή όταν είναι πραγματικά θετικές, ενώ αντίστοιχα ο ρυθμός εσφαλμένων θετικών προβλέψεων (False Positive Rate) εκφράζει τη συχνότητα πρόβλεψης θετικών τιμών από τον ταξινομητή όταν είναι πραγματικά αρνητικές.

$$\text{True Positive Rate(TPR)} = \frac{TP}{TP+FN} \quad (4.4.3)$$

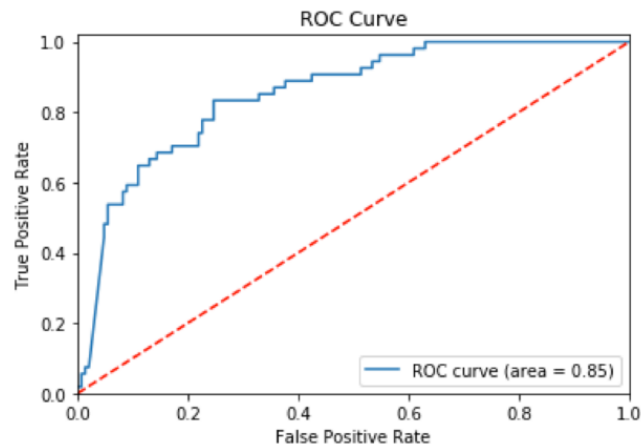
και

$$\text{False Positive Rate(TPR)} = \frac{FP}{TN+FP} \quad (4.4.4)$$

Επιπλέον, ένα μέτρο που υπολογίζει τη συνολική απόδοση του ταξινομητή και λειτουργεί ως δείκτης ποιότητας της ταξινόμησης είναι:

$$\text{Overall Quality(OQ)} = \frac{TP}{TP+FP+FN} \quad (4.4.5)$$

Τέλος, η καμπύλη ROC (ROC-Curve), αποτελεί ένα μέτρο απόδοσης των προβλημάτων ταξινόμησης για διαφορετικές ρυθμίσεις κατωφλίων. Η ROC καμπύλη είναι μία καμπύλη πιθανότητας ενώ το μέτρο AUC (Area Under Curve), που είναι το εμβαδόν της περιοχής κάτω από την καμπύλη ROC, αντιπροσωπεύει το μέτρο ικανότητας διαχωρισμού του ταξινομητή, δηλαδή πόσο καλά μπορεί ο ταξινομητής να διαχωρίζει τις τιμές στις επιμέρους κατηγορίες. Όσο υψηλότερη είναι η τιμή AUC, τόσο καλύτερη η διακριτική ικανότητα του ταξινομητή. Η καμπύλη ROC σχεδιάζεται από τον ρυθμό ορθών θετικών προβλέψεων (TPR) σε σχέση με τον ρυθμό εσφαλμένων θετικών προβλέψεων (FPR), όπου ο πρώτος απεικονίζεται στον άξονα y και ο δεύτερος στον άξονα x αντίστοιχα.



Εικόνα 4.18: Παράδειγμα γραφικής παράστασης καμπύλης ROC.

Κεφάλαιο 5^ο

Σε αυτό το κεφάλαιο αναλύονται οι επιδόσεις και τα αποτελέσματα των δύο μεθοδολογιών ανίχνευσης αερίου, δηλαδή της ανίχνευσης αερίου με χρήση ταξινομητών και της ανίχνευσης αερίου με εξαγωγή υποβάθρου. Αρχικά περιγράφεται ο τρόπος αξιολόγησης των προτεινόμενων μεθοδολογιών, και στη συνέχεια αναλύεται ο τρόπος δημιουργίας των δεδομένων επικύρωσης. Τέλος παρουσιάζονται και αναλύονται τα αποτελέσματα της απόδοσης ανίχνευσης του αερίου και για τις δύο μεθόδους.

5 Αποτελέσματα

5.1 Μεθοδολογία αξιολόγησης αποτελεσμάτων

Στη παρούσα παράγραφο αναλύονται οι διαφορετικοί τύποι σφαλμάτων μέσω δημιουργίας πινάκων σύγχυσης, ενώ επιπρόσθετα παρουσιάζονται και ορισμένα στιγμιότυπα στα οποία γίνεται η ανίχνευση του αερίου. Τα δεδομένα που χρησιμοποιήθηκαν, όπως αναλύθηκαν στην ενότητα [4.1](#), προέρχονται από την καταγραφή δύο βίντεο, ένα στο ορατό και ένα στο θερμικό τμήμα του ηλεκτρομαγνητικού φάσματος. Όπως αναφέρθηκε και στο [4^ο κεφάλαιο](#), στα πλαίσια της διπλωματικής αυτής εφαρμόστηκαν και αξιολογήθηκαν στη μεθοδολογία ταξινομητών δύο είδη ταξινομητών, ενώ στη μεθοδολογία σύνθετης εξαγωγής υποβάθρου ένα μοντέλο από τον συγχρονισμό και την αποκατάστασης γεωμετρίας σκηνής από δύο βίντεο (ένα στο ορατό και ένα στο θερμικό φάσμα). Από τους παράγοντες μετρικής, αναλύθηκαν τα παρακάτω χρησιμοποιούμενα μέτρα για την αξιολόγηση των δύο μεθοδολογιών:

- Πληρότητα ή Ανάκληση (Completeness or Recall)
- Ορθότητα ή Ακρίβεια (Correctness or Precision)
- Συνολική Ποιότητα (Overall Quality)
- Περιοχή κάτω από την καμπύλη ROC (Area Under ROC Curve)

Για την αξιολόγηση της απόδοσης των δύο μεθοδολογιών ανίχνευσης δημιουργήθηκε ένα σύνολο δεδομένων επικύρωσης. Πιο αναλυτικά, επιλέχθηκαν εκατό στιγμιότυπα από το θερμικό βίντεο ανά τριακόσια στιγμιότυπα από τα οποία ψηφιοποιήθηκαν μέσω του ελεύθερου λογισμικού QGIS οι περιοχές της εικόνας στις οποίες ήταν ορατό το αέριο. Στη συνέχεια οι περιοχές που ήταν ορατό το αέριο χαρακτηρίστηκαν ως θετικές με τιμή ένα, ενώ οι υπόλοιπες περιοχές ως αρνητικές με την τιμή μηδέν και τέλος μετατράπηκαν οι ψηφιοποιημένες εικόνες σε μορφή raster. Τα τελικά προϊόντα της ψηφιοποίησης αποτέλεσαν τα δεδομένα επικύρωσης της απόδοσης των δύο μεθοδολογιών.



Εικόνα 5.1: Στιγμιότυπο από το θερμικό βίντεο (αριστερά) και το αποτέλεσμα της ψηφιοποίησης του αερίου για την παραγωγή δεδομένων επικύρωσης (δεξιά).

Στους πίνακες σύγκρισης που δημιουργήθηκαν για την πρώτη μέθοδο, υπολογίστηκαν τα τέσσερα μέτρα του πίνακα, δηλαδή ο συνολικός αριθμός των πραγματικών θετικών, των πραγματικών αρνητικών, των εσφαλμένων θετικών και των εσφαλμένων αρνητικών, αφού πρώτα ετμήθησαν τα προϊόντα της ψηφιοποίησης τύπου raster σε κελιά όπως ακριβώς και κατά τη διαδικασία της ανίχνευσης που αναλύθηκε στην ενότητα [4.2.4](#). Στη συνέχεια, τα μέτρα μετρίτικης υπολογίστηκαν με τη σύγκριση των κελιών από τα προϊόντα ψηφιοποίησης με τα αντίστοιχα κελιά από τα στιγμιότυπα του βίντεο πρόβλεψης. Καθώς τα δεδομένα εκπαίδευσης των ταξινομητών της πρώτης μεθοδολογίας εξάχθηκαν από περιοχές στιγμιότυπων ενός μοναδικού βίντεο και αφορούσαν σε μία σκηνή ενός συγκεκριμένου υποβάθρου, αναμένεται η απόδοση των εκπαιδευόμενων ταξινομητών να διαφοροποιείται σε ετερογενή εξεταζόμενα δεδομένα εισόδου – βίντεο. Σχετικά με την αξιολόγηση της δεύτερης μεθοδολογίας το μέγεθος των μέτρων του πίνακα σύγκρισης αφορούσε στον συνολικό αριθμό των ορθών, και ξανά αντίστοιχα των λανθασμένων, εικονοστοιχείων πρόβλεψης από τη σύγκριση των ανάλογων στιγμιότυπων του βίντεο πρόβλεψης και των ψηφιοποιημένων raster. Αξίζει να σημειωθεί πως στο θερμικό βίντεο που εμφανίζεται το αέριο, οι θερμοκρασιακές διαφορές ποικίλουν, καθώς τα δεδομένα δεν δημιουργήθηκαν σε ελεγχόμενο εργαστηριακό χώρο, αλλά σε εξωτερική τοποθεσία. Συνεπώς, στα στοιχεία του βίντεο εκτός από το ζητούμενο αντικείμενο που είναι το αέριο, στη σκηνή υπάρχουν διερχόμενοι πεζοί, οχήματα και μέτριας έντασης άνεμος που γίνεται αντιληπτός από τα κινούμενα φυλλώματα δέντρων, και από την απότομη αλλαγή κατεύθυνσης του αερίου κ.λπ., παράγοντες που συνετέλεσαν στην δυσκολία ανίχνευσης και απομόνωσης του αερίου. Τέλος, για την ευκολία ανάγνωσης των αποτελεσμάτων, τα παραπάνω μέτρα παρουσιάζονται στις παρακάτω ενότητες με τις ακόλουθες συντομογραφίες:

- Completeness – COM
- Correctness – COR
- Overall Quality – OQ
- Area under ROC curve – AUC

5.2 Δημιουργία δεδομένων επικύρωσης

Όπως συνοπτικά αναφέρθηκε στην ενότητα [5.1](#), τα δεδομένα επικύρωσης της απόδοσης των δύο μεθοδολογιών ανίχνευσης δημιουργήθηκαν με τη βοήθεια του ελεύθερου λογισμικού QGIS, ενώ επιπλέον δημιουργήθηκαν δύο διαφορετικά σετ δεδομένων επικύρωσης για την αξιολόγηση της κάθε τεχνικής. Κατά τη δημιουργία των συνόλων δεδομένων επικύρωσης, ακολούθησε ψηφιοποίηση τόσο των στιγμιότυπων όπου εμφανιζόταν το αέριο, όσο και των στιγμιότυπων που εμφανίζεται μόνο το υπόβαθρο στη σκηνή. Καθώς η μεθοδολογία εντοπισμού αερίου με τη χρήση ταξινομητών είχε ως σκοπό την ένδειξη ύπαρξης ή μη του

αερίου σε συγκεκριμένο κομμάτι της εικόνας του βίντεο δημιουργώντας ένα τετράγωνο κελί γύρω από την περιοχή προβλεπόμενης ύπαρξης του, τα δεδομένα επικύρωσης έπρεπε να μετατραπούν αντίστοιχα από ψηφιοποιημένες περιοχές εικονοστοιχείων πραγματικής ύπαρξης του αερίου σε κελιά- υποπεριοχών της εικόνας που υπήρχε το αέριο. Επίσης στο τελικό βίντεο προβλέπονταν η ύπαρξη αερίου στα κελιά υποπεριοχές της εικόνας, μετά την ολοκλήρωση της ψηφιοποίησης των ground truth εικονοστοιχείων και της μετατροπής των εικόνων αυτών σε μορφή raster, προγραμματιστικά μετατράπηκαν τα πραγματικά θετικά εικονοστοιχεία, σε θετικά πραγματικά κελιά ύπαρξης του αερίου. Η μετατροπή αυτή επιτεύχθηκε με την ανίχνευση των κελιών των raster επικύρωσης όπου η μέση τιμή τους ήταν διάφορη του μηδενός και στη συνέχεια μετατροπή τους σε κελιά μοναδιαίων πινάκων 24x24 εικονοστοιχείων, όπως παρουσιάζεται και στην Εικόνα 5.2:



Εικόνα 5.2: Το αποτέλεσμα ψηφιοποίησης (αριστερά) και το αποτέλεσμα της επεξεργασίας σε θετικά πραγματικά κελιά (δεξιά).

Η δεύτερη μεθοδολογία αφορούσε στην εξαγωγή υποβάθρου από τα βίντεο δύο καμερών διαφορετικού φάσματος και συνεπώς το τελικό βίντεο πρόβλεψης αφορούσε τον εντοπισμό των εικονοστοιχείων του αερίου στο θερμικό βίντεο. Συνεπώς, αυτούσια τα αποτελέσματα της ψηφιοποίησης πραγματικών θετικών τιμών χωρίς περαιτέρω επεξεργασία χρησιμοποιήθηκαν ως δεδομένα επικύρωσης.

5.3 Αποτελέσματα ανίχνευσης αερίου με χρήση ταξινομητών

Κατά την ανάλυση των αποτελεσμάτων ανίχνευσης με τη χρήση ταξινομητών, παρουσιάζονται οι αποδόσεις των δύο ταξινομητών για τους πέντε διαφορετικούς περιγραφείς. Οι παράμετροι κατά τον σχηματισμό των περιγραφέων ήταν σταθεροί, προκειμένου να είναι συγκρίσιμη η απόδοσή τους. Σχετικά με τον περιγραφέα προσανατολισμού ακμών (EOH), έπειτα από δοκιμές διαφορετικών βημάτων (bins) επιλέχθηκε ο αριθμός των 18 βημάτων που επέστρεφε τα βέλτιστα αποτελέσματα, καθώς βήματα μικρότερα της τάξης των 9 ή μεγαλύτερα της τάξης των 36, ήταν είτε πολύ απλοποιημένα είτε πολύ περίπλοκα αντίστοιχα. Σε ορισμένες περιπτώσεις οι διακριτές ικανότητες των διαφορετικών περιγραφέων χαρακτηριστικών, μπορούν να συνδυαστούν προκειμένου να αυξηθεί η διαζευκτική ικανότητα και η ισχύς του ταξινομητή [56], όπως στην περίπτωση εφαρμογής του περιγραφέα HOGHOF. Ο περιγραφέας ιστογραμμάτων προσανατολισμένων διαβαθμίσεων (HOG) αποτελεί μια προέκταση του του περιγραφέα ιστογράμματος προσανατολισμού ακμών (EOH), που σε αντίθεση με τον EOH που υπολογίζει ολόκληρη την εικόνα, ο HOG συλλαμβάνει δομές και εξάγει ιστογράμματα προσανατολισμού ακμών από ένα πλέγμα της εικόνας, με τη βοήθεια ενός παραθύρου ανίχνευσης προκαθορισμένου μεγέθους.

Πίνακας 5-1: Αποτελέσματα απόδοσης του ταξινομητή SVM για διαφορετικούς περιγραφείς χαρακτηριστικών.

Ταξινομητής SVM					
Περιγραφέας	Μήκος	COM(%)	COR(%)	OQ(%)	AUC
<i>F(EOH)</i>	18	35.36	77.21	32.06	0.72
<i>F(LBP^{u2})</i>	59	22.47	68.28	20.35	0.63
<i>F(LBP^{ru2})</i>	10	23.53	73.45	26.23	0.56
<i>F(LBPTOP)</i>	20	31.74	70.45	28.01	0.69
<i>F(HOGHOF)</i>	177	33.97	81.17	31.49	0.71

Συγκεντρωτικά τα αποτελέσματα απόδοσης του Πίνακα 5-1 της ταξινόμησης με τον SVM, ο τελεστής LBP Uniform παρουσιάζει τα χαμηλότερα ποσοστά απόδοσης από όλους του περιγραφείς με ποσοστό 22.47% πληρότητα, 68.28% ορθότητα και συνολική ποιότητα 20.35%. Επιπρόσθετα, το μεγαλύτερο ποσοστό ορθότητας ή αλλιώς ακρίβειας είχε η ταξινόμηση του περιγραφέα HOGHOF, με συνολικό ποσοστό ορθών ταξινομήσεων και των δύο κατηγοριών, ύπαρξης ή μη αερίου, 81.17%. Αντίστοιχα, η ταξινόμηση του περιγραφέα EOH από τον ταξινομητή SVM συνολικά επέστρεψε τα υψηλότερα ποσοστά απόδοσης, πληρότητας με ποσοστό 35.36% , συνολικής ποιότητας της ταξινόμησης 32.06% και περιοχής κάτω από την καμπύλη ROC 0.72, με μέγιστη τιμή το 1.

Πίνακας 5-2: Αποτελέσματα απόδοσης του ταξινομητή AdaBoost για διαφορετικούς περιγραφείς χαρακτηριστικών.

Ταξινομητής AdaBoost					
Περιγραφέας	Μήκος	COM(%)	COR(%)	OQ(%)	AUC
<i>F(EOH)</i>	18	35.14	77.10	32.82	0.72
<i>F(LBP^{u2})</i>	59	36.85	72.69	32.37	0.73
<i>F(LBP^{ru2})</i>	10	37.40	79.50	34.27	0.74
<i>F(LBPTOP)</i>	20	31.74	77.40	29.04	0.69
<i>F(HOGHOF)</i>	177	35.04	81.23	32.41	0.71

Από τη ταξινόμηση περιγραφέων με τον AdaBoost, όπως φαίνεται στον Πίνακα 5-2, η απόδοση της ταξινόμησης του περιγραφέα ομοιόμορφων ανεξάρτητων περιστροφής προτύπων ήταν από τις βέλτιστες και ξανά υψηλότερη σε σύγκριση με την ταξινόμηση του LBP^{ru2}, με ποσοστό πληρότητας 37.40%, ορθότητας 79.50% , συνολικής ποιότητας 34.27% και τιμής AUC 0.73. Το υψηλότερο ποσοστό πληρότητας σημειώθηκε από την ταξινόμηση του περιγραφέα

HOGHOF με 81.23%, ποσοστό υψηλότερο συγκριτικά και με τα ποσοστά απόδοσης της ταξινόμησης με τον SVM. Η χαμηλότερη απόδοση ταξινόμησης με τον ταξινομητή AdaBoost σημειώθηκε με την ταξινόμηση του περιγραφέα LBPTOP με ποσοστό πληρότητας 31.47%, συνολική ποιότητα 29.04% και τιμή AUC 0.69. Με γνώμονα την αύξηση της απόδοσης της τεχνικής ανίχνευσης αερίου με τη χρήση ταξινομητών, αξιολογήθηκε ο συνδυασμός των εκπαιδευόμενων ταξινομητών. Θεωρώντας σταθερό το είδος του ταξινομητή (SVM και AdaBoost), υλοποιήθηκαν τρεις εφαρμογές συνδυασμού των ταξινομητών μεταξύ τους. Πιο συγκεκριμένα, δημιουργήθηκαν τρία μοντέλα πρόβλεψης με βάση τη φύση των περιγραφών χαρακτηριστικών από τους οποίους εκπαιδεύτηκαν οι ταξινομητές, με αποτέλεσμα να γίνει μία εφαρμογή πρόβλεψης από "στατικούς" ταξινομητές (F(EOH), F(LBP^{u2}) και F(LBPⁿⁱ²)), μία από "δυναμικούς" ταξινομητές (F(LBPTOP) και F(HOGHOF)) και τέλος μία από τον συνδυασμό των στατικών και των δυναμικών ταξινομητών και για τα δύο διαφορετικά είδη ταξινομητών.

Πίνακας 5-3: Αποτελέσματα απόδοσης ταξινόμησης από τον συνδυασμό εκπαιδευόμενων ταξινομητών διαφορετικών περιγραφών χαρακτηριστικών.

Περιγραφείς χαρακτηριστικών	Ταξινομητής SVM			
	COM(%)	COR(%)	OQ(%)	AUC
Στατικοί	19.38	72.51	18.05	0.62
Δυναμικοί	23.53	86.33	22.69	0.65
Συνδυασμός	36.32	78.21	32.98	0.72
	Ταξινομητής AdaBoost			
Στατικοί	26.30	81.52	24.82	0.67
Δυναμικοί	25.03	89.70	24.33	0.65
Συνδυασμός	38.13	82.11	35.20	0.73

Στον Πίνακα 5-3 παρουσιάζονται τα αποτελέσματα απόδοσης ταξινόμησης έπειτα από το συνδυασμό ταξινομητών στατικών περιγραφών χαρακτηριστικών και αντίστοιχα, δυναμικών περιγραφών χαρακτηριστικών. Όπως γίνεται αντιληπτό από τη σύγκριση των αποτελεσμάτων των Πινάκων 5-1, 5-2 και 5-3, ο συνδυασμός ταξινομητών επέφερε τα υψηλότερα ποσοστά απόδοσης ταξινόμησης και των δύο κατηγοριών. Πιο συγκεκριμένα, από τον συνδυασμό ταξινομητών SVM, τα υψηλότερα ποσοστά απόδοσης ταξινόμησης, όσον αφορά στην πληρότητα με ποσοστό 36.32%, την συνολική ποιότητα με 32.98% και τιμή AUC με 0.72, παρουσίασε ο συνδυασμός στατικών και δυναμικών ταξινομητών, ενώ η ταξινόμηση με δυναμικούς ταξινομητές είχε το υψηλότερο ποσοστό ορθότητας με 86.33%. Αντίθετα, η ταξινόμηση με συνδυασμό μόνο των στατικών περιγραφών χαρακτηριστικών είχε τα χαμηλότερα ποσοστά απόδοσης. Επιπρόσθετα, η ταξινόμηση με συνδυασμό ταξινομητών AdaBoost επέστρεψε πανομοιότυπα ποσοστά απόδοσης, με το συνδυασμό των στατικών και δυναμικών να εμφανίζει ξανά τα υψηλότερα ποσοστά με 38.13% πληρότητα, 35.20% συνολική

ποιότητα και 0.73 τιμή AUC, ενώ το μεγαλύτερο ποσοστό ορθότητας είχε πάλι η ταξινόμηση με τους δυναμικούς ταξινομητές με 89.70%.



Εικόνα 5.3: Επιτυχημένες ανιχνεύσεις με χρήση AdaBoost για την ταξινόμηση LBP^{ru2} χαρακτηριστικών.

Η χαμηλή απόδοση ταξινόμησης του τελεστή LBP Uniform (LBP^{u2}) σε σχέση με τον ομοιόμορφο ανεξάρτητου περιστροφής LBP (LBP^{ru2}), ο οποίος σημείωσε ελαφρώς καλύτερα αποτελέσματα απόδοσης από την ταξινόμηση με τον SVM, με 23.53% πληρότητα, 73.45% ορθότητα και 26.23% συνολική ποιότητα, σχετίζεται πιθανότατα με το μικρότερο μήκος ιστόγραμμα του δεύτερου, που φαίνεται να είναι αρκετό για να συλλάβει την πολυπλοκότητα της υφής στα στιγμιότυπα με αέριο.



Εικόνα 5.4: Επιτυχημένες ανιχνεύσεις με χρήση SVM για την ταξινόμηση EOH χαρακτηριστικών (αριστερά) και συνδυασμού στατικών και δυναμικών SVM ταξινομητών (δεξιά).

Επιπλέον, η καλή διακριτική ικανότητα του περιγραφέα προσανατολισμού ακμών, και από τις δύο ταξινομήσεις, μπορεί να αποδοθεί στο γεγονός ότι ορισμένες περιοχές χωρίς αέριο της εικόνας παρουσιάζουν υψηλά μεγέθη ακμών, καθώς υπάρχουν κτήρια, αμάξια στη σκηνή, και σε συνδυασμό με τις γρήγορες μεταβολές στον προσανατολισμό, από κινούμενα οχήματα, οι εξαγόμενες πληροφορίες από τον περιγραφέα αυτόν από μόνες τους μπορούσαν να καθορίσουν ικανοποιητικά την ύπαρξη ή μη αερίου στα υποψήφια στιγμιότυπα. Όσον αφορά τον χωροχρονικό περιγραφέα LBPTOP, τα χαμηλά ποσοστά απόδοσης στην ταξινόμηση και στα δύο είδη ταξινομητών ίσως να οφείλονται στο γεγονός ότι σε μια ταχεία κίνηση του αερίου, ο περιγραφέας να χάνει το ίχνος του.



Εικόνα 5.5: Επιτυχημένες ανιχνεύσεις με χρήση συνδυασμού στατικών και δυναμικών AdaBoost ταξινομητών (αριστερά) και συνδυασμού μόνο δυναμικών AdaBoost ταξινομητών (δεξιά).

Ο περιγραφέας αυτός μπορεί να βελτιωθεί εάν αυξηθούν τα κυβοειδή δεδομένα εκπαίδευσης, καθώς ακόμα και αν από τη θεωρία καθορίζεται πως η απόδοση ενός ταξινομητή SVM δεν σχετίζεται με το μέγεθος των δεδομένων εκπαίδευσης αλλά από το πόσο αντιπροσωπευτικά είναι τα δείγματα, στην πράξη μετά από επαναληπτικές πειραματικές διαδικασίες, έγινε αντιληπτό πως αυξάνοντας το μέγεθος των δεδομένων εκπαίδευσης και προσθέτοντας μεγαλύτερη ποικιλία στα δείγματα παράγονται περισσότερα διανύσματα υποστήριξης, βελτιώνοντας ελαφρώς την απόδοση του ταξινομητή.



Εικόνα 5.6: Εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού ταξινομητών SVM (αριστερά) και συνδυασμού ταξινομητών AdaBoost (δεξιά).

Επιπλέον, άλλη μία αιτία που μπορεί να δικαιολογήσει τα χαμηλά ποσοστά απόδοσης ταξινόμησης του LBPTOP αλλά και του HOGHOF, οφείλεται στην ίδια την κίνηση που προσπαθούν να περιγράψουν, καθώς η διάχυση των αερίων είναι αρκετά περίπλοκη και είναι αρκετά δύσκολο να προσδιορισθεί ποσοτικά. Ταυτόχρονα, συγκρίνοντας τα δύο είδη ταξινομητών είναι εμφανές πως ο ταξινομητής Adaboost παρουσιάζει σχετικά καλύτερα αποτελέσματα ταξινόμησης σχεδόν από όλους τους περιγραφείς. Η υπεροχή στην απόδοση από τον Adaboost ίσως σχετίζεται με το γεγονός ότι έχει την ικανότητα να "μάθει" μη γραμμικά όρια απόφασης, χαρακτηριστικό το οποίο είναι χρήσιμο στην περίπτωση που τα δεδομένα εκπαίδευσης δεν μπορούν να διαχωριστούν γραμμικά.



Εικόνα 5.7: Επιτυχείς και εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού στατικών ταξινομητών AdaBoost (αριστερά) και SVM (δεξιά).

Τέλος, τα χαμηλά ποσοστά, δηλαδή των τιμών κάτω του 50%, που σημειώνονται στα μέτρα της πληρότητας και της συνολικής ακρίβειας, ίσως σχετίζονται με το γεγονός ότι τα δεδομένα εκπαίδευσης των ταξινομητών ήταν κομμάτια εικόνας ενός μοναδικού βίντεο. Το μικρό δείγμα εκπαίδευσης δεν επέτρεψε την κατασκευή ισχυρών ταξινομητών, ενώ ταυτόχρονα η κατασκευή περιγραφέων χαρακτηριστικών μιας τέτοιας πολύπλοκης φύσης αντικειμένου, όπως το αέριο, περιέπλεξε ακόμα περισσότερο την εκπαίδευση ταξινομητών με υψηλή διακριτική ικανότητα.



Εικόνα 5.8: Επιτυχείς και εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού δυναμικών ταξινομητών AdaBoost (αριστερά) και SVM (δεξιά).

5.4 Αποτελέσματα απόδοσης ανίχνευσης αερίου με εξαγωγή υποβάθρου

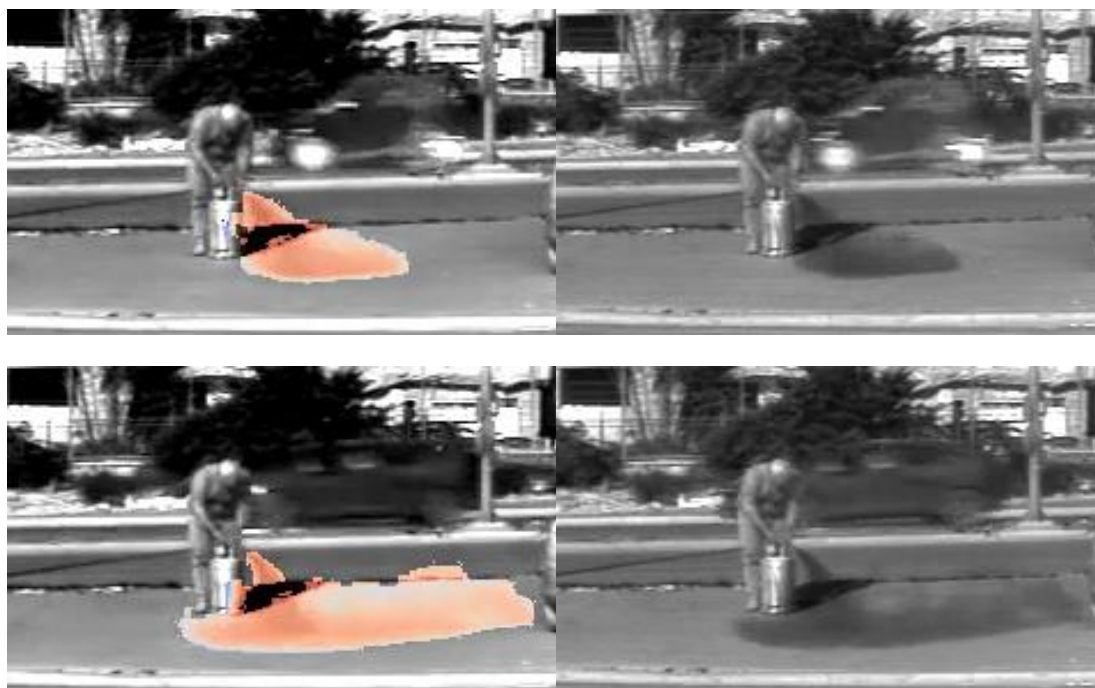
Η δεύτερη μέθοδος ανίχνευσης του αερίου βασιζόταν στον εντοπισμό και εξαγωγή του υποβάθρου στις θερμικές εικόνες με σκοπό την οριοθέτηση μονάχα των εικονοστοιχείων του αερίου. Όπως αναφέρθηκε και στην ενότητα [5.1](#), η επικύρωση της μεθόδου αυτής σχετίζεται με την ορθή ανίχνευση των εικονοστοιχείων που ανήκουν στο αέριο, και όχι τμημάτων της εικόνας με αέριο που ακολουθήθηκε στην επικύρωση της μεθόδου ανίχνευσης με χρήση ταξινομητών. Τα παραγόμενα raster από την ψηφιοποίηση των δεδομένων επικύρωσης χρησιμοποιήθηκαν αυτούσια χωρίς την περαιτέρω επεξεργασία τους. Συνεπώς η σύγκριση των δεδομένων πρόβλεψης και των δεδομένων επικύρωσης αφορούσε στη σύγκριση των τιμών των εικονοστοιχείων. Αφού υπολογίσθηκαν αρχικά οι τελεστές πραγματικών θετικών, πραγματικών αρνητικών, εσφαλμένων θετικών και εσφαλμένων αρνητικών για τα στιγμιότυπα

πρόβλεψης από τα αντίστοιχα στιγμιότυπα επικύρωσης, εκτιμήθηκαν οι μέσοι όροι για κάθε τελεστή για τον υπολογισμό των ποσοστών πληρότητας, ορθότητας και συνολικής ποιότητας.

Πίνακας 5-4: Αποτελέσματα απόδοσης της κατηγοριοποίησης των εικόνων από τη μέθοδο εξαγωγής υποβάθρου.

Μέθοδος Εξαγωγής Υποβάθρου				
Απόδοση	COM(%)	COR(%)	OQ(%)	AUC
	88.59	55.40	51.71	0.93

Από τον Πίνακα 5-4 γίνεται αντιληπτό ότι τα αποτελέσματα απόδοσης της μεθόδου εξαγωγής υποβάθρου επέστρεψαν υψηλότερα ποσοστά απόδοσης ταξινόμησης. Πιο αναλυτικά, η μέθοδος εξαγωγής υποβάθρου επέστρεψε ποσοστό πληρότητας 88.59%, που αντιπροσωπεύει το γεγονός ότι η πλειοψηφία των εικονοστοιχείων του αερίου κατηγοριοποιήθηκαν σωστά, ποσοστό της συνολικής ποιότητας 51.71% και AUC σχεδόν 1 με τιμή 0.93.



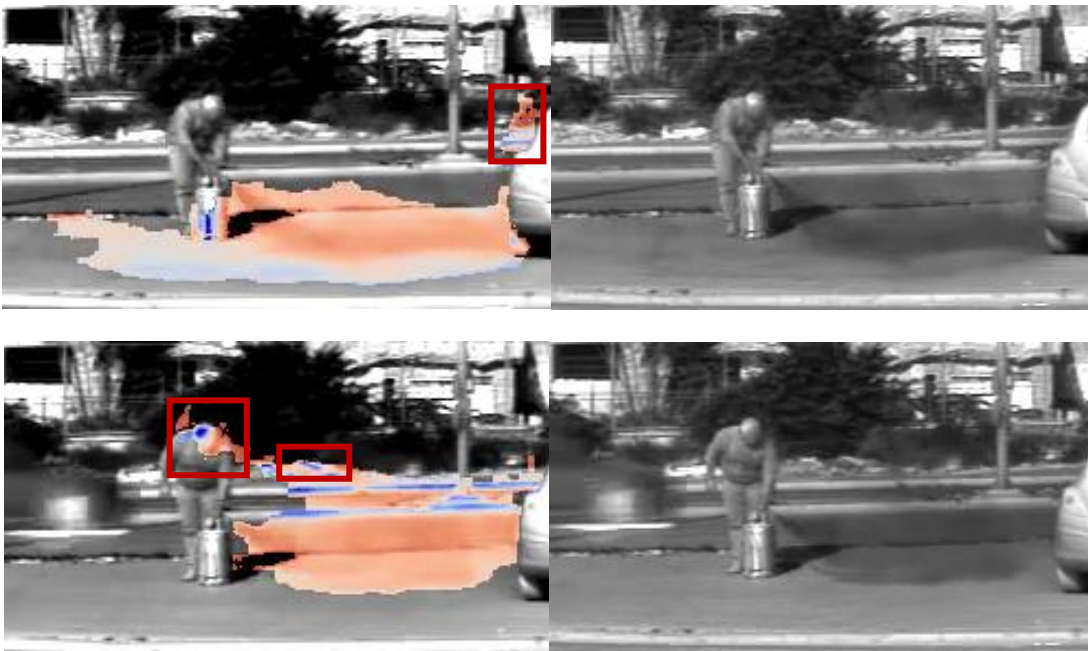
Εικόνα 5.9: Επιτυχείς ανιχνεύσεις εικονοστοιχείων αερίου με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).

Αντιθέτως το ποσοστό της ορθότητας, που εκφράζει την ορθή πρόβλεψη και των δύο κατηγοριών, ελαττώθηκε σχεδόν κατά 25% από τα αντίστοιχα αποτελέσματα της μεθόδου των ταξινομητών, με ποσοστό 55.40%. Επιπρόσθετα παρατηρείται ότι το ποσοστό της συνολικής ποιότητας της μεθόδου εξαγωγής υποβάθρου ξεπέρασε τα αντίστοιχα ποσοστά της μεθόδου των ταξινομητών σχεδόν κατά 15%.



Εικόνα 5.10: Εσφαλμένες ανιχνεύσεις εικονοστοιχείων αερίου (με κόκκινο περίγραμμα) με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).

Συμπερασματικά, καθώς το μέτρο της ορθότητας-ακρίβειας σχετίζεται με την ορθή πρόβλεψη και των δύο κατηγοριών, υπάρχει η πιθανότητα αυτή η μείωση του ποσοστού να οφείλεται σε ελλιπή ψηφιοποίηση των δεδομένων επικύρωσης. Η ύπαρξη του αερίου στην εικόνα του βίντεο δεν είναι πάντα παρατηρήσιμη, αφού οι τιμές των εικονοστοιχείων του υποβάθρου, στην ύπαρξη αερίου, μπορούν να έχουν μια πολύ μικρή μεταβολή στις τιμές τους, γεγονός που μπορεί να ανιχνευθεί από τις αλγοριθμικές διαδικασίες εξαγωγής υποβάθρου, αλλά όχι με γυμνό μάτι κατά τη διαδικασία της ψηφιοποίησης των δεδομένων επικύρωσης.



Εικόνα 5.11: Επιτυχείς και εσφαλμένες (με κόκκινο περίγραμμα) ανιχνεύσεις εικονοστοιχείων αερίου με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).

Κεφάλαιο 6^ο

Το κεφάλαιο αυτό ολοκληρώνει την εργασία παρέχοντας μια σύνοψη περίληψη των αποτελεσμάτων και των συμπερασμάτων στα οποία κατέληξε. Επιπρόσθετα, προτείνονται τρόποι ανάπτυξης της παρούσας διπλωματικής στο τομέα της αυτόματης ανίχνευσης αερίων.

6 Συμπεράσματα

6.1 Ειδικά και Γενικά Συμπεράσματα

Η παρούσα διπλωματική, παρά την πολυπλοκότητα του προβλήματος της αυτόματης ανίχνευσης αερίων, με τη χρήση αλγοριθμικών διαδικασιών υπολογιστικής μάθησης και υπολογιστικής όρασης κατάφερε να παρουσιάσει και να ολοκληρώσει με επιτυχία δύο διαφορετικές προσεγγίσεις ανίχνευσης. Όπως έχει προαναφερθεί, η φύση του προβλήματος είναι σύνθετη από μόνη της λόγω των ευμετάβλητων χαρακτηριστικών του αερίου, ενώ οι περισσότερες από τις εφαρμογές ανίχνευσης που εξετάστηκαν κατά την βιβλιογραφική ανάλυση, αφορούσαν σε πειραματικές διατάξεις σε σταθερό και ελεγχόμενο περιβάλλον, ή χρήση ανοικτών δεδομένων από πλατφόρμες στο διαδίκτυο. Αντίθετα, η παρούσα εργασία υλοποιήθηκε βασισόμενη σε μια εξωτερική πειραματική διάταξη όπου εκτός από τη δυσκολία ανίχνευσης του αερίου λόγω της φύσης του, το υπόβαθρο της σκηνής ήταν γεμάτο από κινούμενα αντικείμενα και από μεταβαλλόμενες περιβαλλοντικές παραμέτρους, όπως οι διαφορετικές κατευθύνσεις του αέρα. Τελικά, και οι δύο προσεγγίσεις που προτείνονται από την παρούσα εργασία πέτυχαν να ανιχνεύσουν και να οπτικοποιήσουν σε έναν πολύ ικανοποιητικό βαθμό το αέριο προπάνιο.

Επιπλέον, και οι δύο τεχνικές ήταν σε θέση να εντοπίσουν το αέριο σε ακραίες συνθήκες διαφάνειας και κίνησης αντικειμένων μέσα από το νέφος του αερίου. Καθώς η ανίχνευση του κάθε αερίου σχετίζεται άμεσα με τα ατμοσφαιρικά παράθυρα απορρόφησης της ακτινοβολίας που το χαρακτηρίζει, οι ίδιες τεχνικές θα μπορούσαν να εφαρμοστούν για την ανίχνευση οποιοδήποτε αερίου που μπορεί να συλληφθεί από την αντίστοιχη θερμική κάμερα. Ταυτόχρονα εάν ζητούμενο του προβλήματος είναι η έγκαιρη ειδοποίηση εντοπισμού διαρροής αερίου, οι παραπάνω προτεινόμενες μεθοδολογίες παρουσιάζουν ικανοποιητικά αποτελέσματα.

Στη παρούσα διπλωματική εργασία υλοποιήθηκαν δύο μεθοδολογίες εφαρμογής αυτόματης ανίχνευσης αερίων, όπου η μία βασίζεται στις μεθόδους της υπολογιστικής μάθησης και υπολογιστικής όρασης, ενώ η δεύτερη στον συνδυασμό βίντεο από κάμερες διαφορετικού φάσματος για την εξαγωγή του υποβάθρου της σκηνής για την απομόνωση του αερίου. Η μεθοδολογία της υπολογιστικής μάθησης, έδωσε ικανοποιητικά ποσοστά ανίχνευσης του αερίου, ανάλογα με χρήση του εκάστοτε περιγραφέα, ενώ η μέθοδος εξαγωγής υποβάθρου επέτυχε τα υψηλότερα ποσοστά ορθής ανίχνευσης του αερίου και ταυτόχρονα τα χαμηλότερα ποσοστά εσφαλμένης ανίχνευσης.

Οι δύο τεχνικές αναπτύχθηκαν με βάση τα στιγμιότυπα βίντεο που λήφθηκαν από δύο στατικές κάμερες, της θερμικής κάμερας Ventus™ της εταιρείας SIERRA OLYMPIC και μίας συμβατικής κάμερας του ορατού φάσματος. Η πειραματική διάταξη ήταν σε εξωτερικό χώρο, με αποτέλεσμα η σκηνή να ήταν συνεχώς μεταβαλλόμενη, ενώ τα αέρια διαταρασσόταν μη ομοιόμορφα από τον άνεμο. Κατά τη μεθοδολογία της υπολογιστικής μάθησης τα στιγμιότυπα του βίντεο διαιρέθηκαν σε μικρά τμήματα (block) στα οποία ανιχνευόντουσαν τα κινούμενα αντικείμενα διαμέσου της τεχνικής αφαίρεσης υποβάθρου μέσω της τεχνικής του επιλεκτικά μεταβαλλόμενου μοντέλου. Στα τμήματα της εικόνας που ανιχνευόταν κίνηση αντικειμένων, εξαγόntonταν στατικά και δυναμικά χαρακτηριστικά περιγραφών τα οποία στη συνέχεια ταξινομήθηκαν για τον προσδιορισμό της ύπαρξης ή μη αερίου.

Πιο αναλυτικά, χρησιμοποιήθηκαν στατικοί περιγραφείς χαρακτηριστικών, όπως ο περιγραφέας τοπικών δυαδικών προτύπων (LBP) για την ανάλυση της υφής της εικόνας, ο περιγραφέας ιστογράμματος προσανατολισμού ακμών (EOH) και κλίσεων διαβαθμίσεων (HOG) για την ανάλυση του περιγράμματος των αντικειμένων, αλλά και χωροχρονικοί περιγραφείς ιστογραμμάτων οπτικής ροής (HOF) και τοπικών δυαδικών προτύπων τριών διαστάσεων (LBPTOP). Τα εξαγόμενα χαρακτηριστικά μεταβιβάζονται σε έναν ταξινομητή υποστήριξης διανυσμάτων (SVM) και σε έναν ταξινομητή AdaBoost για τη σύγκριση της απόδοσης της ταξινόμησης όπου βρέθηκαν σημαντικές διαφορές. Κατά την ταξινόμηση των τμημάτων των στιγμιότυπων, τμήματα που παρουσίαζαν κίνηση μετά την αφαίρεση υποβάθρου, μόνο για ένα ή δύο συνεχόμενα στιγμιότυπα αφαιρέθηκαν από τα τμήματα ενδιαφέροντος για την ταξινόμηση τους.

Η δεύτερη μεθοδολογία, αναπτύχθηκε με βάση τα στιγμιότυπα από δύο βίντεο που λήφθηκαν από διαφορετικές ζώνες του φάσματος της ακτινοβολίας, ενός από το μέσο υπέρυθρο και ενός από το ορατό φάσμα. Ο συγχρονισμός των δύο βίντεο, με την ταυτόχρονη αποκατάσταση της γεωμετρίας της σκηνής, έγινε με γνώμονα την βέλτιστη εξαγωγή του υποβάθρου, με τον συνδυασμό δύο τεχνικών αφαίρεσης, της αφαίρεσης υποβάθρου μέσω τεχνικής διαδοχικών στιγμιότυπων και της αφαίρεσης υποβάθρου μέσω τεχνικής ανάπτυξης επιλεκτικά μεταβαλλόμενου μοντέλου, ενώ στη συνέχεια υλοποιήθηκαν μία σειρά από αλγοριθμικές διαδικασίες για την απομόνωση των εικονοστοιχείων που ανήκουν μόνο στο αέριο. Οι δύο τεχνικές αυτόματης ανίχνευσης αερίου, αναπτύχθηκαν σε γλώσσα Python 2.7, με τη τεχνική της υπολογιστικής μάθησης να επιτυγχάνει χρόνους επεξεργασίας 1.5 δευτερολέπτου ανά στιγμιότυπο σε φορητό υπολογιστική χαμηλής κατανάλωσης ενέργειας.

Η μεθοδολογία της ανίχνευσης με τεχνικές υπολογιστικής μάθησης ήταν ικανή να εντοπίσει το αέριο στα στιγμιότυπα του θερμικού βίντεο, με την πιο πετυχημένη ανίχνευση να επιτυγχάνεται από τον συνδυασμό των δυναμικών και των στατικών περιγραφών με ποσοστά 38.13% πληρότητα και 82.11% ορθότητα προβλέψεων από την ταξινόμηση με έναν ταξινομητή AdaBoost, ενώ αντίστοιχα η μεθοδολογία της αφαίρεσης υποβάθρου επέστρεψε 88.59% ποσοστό πληρότητας και 51.71% ποσοστό ορθότητας. Πιο συγκεκριμένα και όσο αφορά την απόδοση των μεθόδων που αναλύθηκαν στο προηγούμενο κεφάλαιο, ο ταξινομητής AdaBoost επέστρεψε υψηλότερα ποσοστά σε σχέση με τις μηχανές διανυσμάτων υποστήριξης, γεγονός που ίσως σχετίζεται στο ότι ο ταξινομητής αυτός έχει την ικανότητα να εκπαιδεύεται από μη γραμμικά όρια απόφασης. Αντίστοιχα, χαρακτηριστικό επιτυχούς εκπαίδευσης του ταξινομητή SVM είναι η αντιπροσωπευτικότητα των δεδομένων εκπαίδευσης, που στην περίπτωση της παρούσας εργασίας, δεν υπήρχε ποικιλία δεδομένων λόγω του ότι δημιουργήθηκαν αποκλειστικά από τα στιγμιότυπα ενός μοναδικού βίντεο.

Από τη σύγκριση των περιγραφών χαρακτηριστικών, οι EOH και LBP^{m2} επέστρεψαν τα υψηλότερα ποσοστά απόδοσης από τους ταξινομητές SVM και AdaBoost αντίστοιχα. Η καλή διακριτική ικανότητα του περιγραφέα προσανατολισμού ακμών ίσως να αποδίδεται στο ότι ορισμένες περιοχές χωρίς αέριο της εικόνας παρουσιάζουν υψηλά μεγέθη ακμών, ενώ οι γρήγορες μεταβολές στον προσανατολισμό, από τα κινούμενα αντικείμενα, επέστρεψαν

καθοριστικές πληροφορίες που μπορούσαν να καθορίσουν ικανοποιητικά την ύπαρξη ή μη αερίου στα υποψήφια στιγμιότυπα. Επιπλέον, ο περιγραφέας LBP^{m2} ο οποίος χαρακτηρίζεται από μικρότερου μήκους ιστόγραμμα, προέκυψε πως να είναι αρκετός για να συλλάβει την πολυπλοκότητα της υφής στα στιγμιότυπα με αέριο.

Αντιθέτως, η δημιουργία και η εξαγωγή χωροχρονικών περιγραφέων δεν οδήγησε σε βελτιωμένη απόδοση ταξινόμησης καθώς δεν ήταν σε θέση να διακρίνουν ικανοποιητικά μεταξύ των αντικειμένων αερίου και του υποβάθρου, ενώ σχετικά με τη απόδοση ταξινόμησης των δύο ταξινομητών, ο ταξινομητής διανυσμάτων υποστήριξης επέστρεψε πιο ικανοποιητικά αποτελέσματα. Επιπλέον, ο αριθμός των ψευδών ανιχνεύσεων, φαίνεται να εξαρτάται σε μεγάλο βαθμό από τον τύπο του περιγραφέα χαρακτηριστικών που χρησιμοποιείται, ενώ τα ποσοστά των εσφαλμένων θετικών ανιχνεύσεων ήταν πολύ χαμηλά από 0.19% έως 1.9%. Επιπρόσθετα, τα ποσοστά των ορθών θετικών ανιχνεύσεων σε ορισμένες περιπτώσεις ήταν πολύ χαμηλά, όπως στην περίπτωση ανίχνευσης αερίου με χρήση LBP^{m2} περιγραφέα και SVM ταξινομητή. Οι περιορισμοί της απόδοσης που περιγράφονται παραπάνω, εν μέρει μπορεί να οφείλονται στο περιορισμένης δεδομένων ποικιλίας στο σύνολο εκπαίδευσης, που δημιουργήθηκε καθώς περιλαμβάνει 7485 στατικά παραδείγματα και 262 δυναμικά κυβοειδή παραδείγματα από περιοχές εικόνας ενός μοναδικού βίντεο με 3800 στιγμιότυπα.

Η προσέγγιση της μηχανικής μάθησης για την ανίχνευση αερίων προσφέρει το πλεονέκτημα ότι είναι σε θέση να επεξεργαστεί ισχυρούς πολυδιάστατους περιγραφικούς δείκτες σε μικρό υπολογιστικό χρόνο καθιστώντας εφικτή την επεξεργασία σε πραγματικό χρόνο. Σε αντίθεση, η δεύτερη μεθοδολογία στηρίζεται στην εξαγωγή υποβάθρου της σκηνής η οποία υλοποιείται σε μεταγενέστερο χρόνο, κάνοντας την ανίχνευση σε πραγματικό χρόνο αδύνατη. Κατά την διάρκεια υλοποίησης των δύο μεθοδολογιών, εντοπίστηκαν επίσης ορισμένες αδυναμίες, κυρίως κατά την ανίχνευση σε περιοχές αερίου όπου οι τιμές της αντίθεσης ήταν ιδιαίτερα χαμηλές.

Επιπλέον, η πολυπλοκότητα της σκηνής με τη μεγάλη ποικιλία των κινούμενων αντικειμένων σε διαφορετικές κατευθύνσεις και η ύπαρξη ανέμου μέτριας έντασης δυσχέρανε την κατασκευή μιας ισχυρής μεθοδολογίας ανίχνευσης και καθιστούσε απαραίτητη την ανάγκη συλλογής και επεξεργασίας πολύ υψηλού μεγέθους δεδομένων κατάρτισης.

6.2 Προτάσεις για την εξέλιξη της παρούσας διπλωματικής

Κατά τη διάρκεια εκπόνησης της παρούσας διπλωματικής πολλά πεδία βρέθηκαν που μπορούν να βελτιωθούν και να αναπτυχθούν περαιτέρω. Λόγω της περίπλοκης φύσης του προβλήματος εντοπισμού αερίων, τα πεδία πιθανής βελτίωσης είναι πολλά από βελτίωση των δεδομένων εκπαίδευσης μέχρι την εφαρμογή μεθόδων μεταγενέστερης επεξεργασίας.

6.2.1 Διερεύνηση ποικιλίας του Συνόλου Δεδομένων

Η επέκταση των διαφορετικών περιπτώσεων στο σύνολο των δεδομένων που χρησιμοποιούνται για την εκπαίδευση φαίνεται να είναι απαραίτητη για την αύξηση των ποσοστών απόδοσης ταξινόμησης και από τους δύο ταξινομητές. Συνεπώς, η επιλογή ενός εκτεταμένου ποικιλόμορφου συνόλου δεδομένων εκπαίδευσης με πολλές διαφορετικές περιπτώσεις θετικών και αρνητικών δειγμάτων κρίνεται απαραίτητη. Στα στιγμιότυπα του

βίντεο που απεικόνιζαν το αέριο και από τα οποία δημιουργήθηκαν δεδομένα εκπαίδευσης, αυτό ανιχνεύθηκε με επιτυχία, ενώ στις διαφορετικές περιπτώσεις που η μορφή του αερίου διαφοροποιούνταν αρκετά από τα δεδομένα εκπαίδευσης, οι ταξινομητές δεν ήταν σε θέση να διακρίνουν την ύπαρξη ή μη. Συνεπώς, ενώ το μέγεθος του συνόλου εκπαίδευσης ήταν μεγάλο, έλειπε η διαφορετικότητα μεταξύ των δειγμάτων. Η βελτίωση των ταξινομητών θα μπορούσε να επιτευχθεί περαιτέρω με τη συλλογή των εσφαλμένων θετικών ανιχνεύσεων και στη συνέχεια την εκ νέου εκπαίδευση των ταξινομητών με τις συγκεκριμένες ανιχνεύσεις ως αρνητικά δεδομένα εκπαίδευσης.

6.2.2 Βελτιώσεις στην εξαγωγή χαρακτηριστικών

Εκτός από την επέκταση του συνόλου δεδομένων εκπαίδευσης, προτείνεται η βελτίωση και περαιτέρω ανάπτυξη των περιγραφών εξαγωγής χαρακτηριστικών, καθώς αυτοί αποτελούν και τα δεδομένα εισόδου εκπαίδευσης των ταξινομητών. Μια διαφορετική προσέγγιση με σκοπό την βελτίωση των περιγραφών θα μπορούσε να είναι η εξαγωγή τους σε πολλές διαφορετικές κλίμακες [55] της εικόνας με μία εφαρμογή παραθύρου κύλισης αντί να εξετάζονται σταθερά κελιά (block) του κάθε στιγμιότυπου [58], [59], [46], [51]. Επιπρόσθετα, θα μπορούσαν χρησιμοποιηθούν περισσότερα κυβοειδή δεδομένα εκπαίδευσης των δυναμικών περιγραφών χαρακτηριστικών κατά την εκπαίδευση πιο ισχυρών ταξινομητών. Στην παρούσα έρευνα, οι δυναμικοί περιγραφείς μπορεί να μην επέστρεψαν την επιθυμητή διακριτική ικανότητα κατά την εκπαίδευση ταξινομητών, παρόλα αυτά η κίνηση είναι ένα σημαντικό χαρακτηριστικό των νεφών αερίων και συνεπώς έχει μεγάλες προοπτικές στην δημιουργία ισχυρών ταξινομητών. Μία πιθανή μέθοδος χρήσης του χαρακτηριστικού της κίνησης είναι για παράδειγμα η μοντελοποίηση της δυναμικής υφής, που παρουσιάζει το αέριο, με τη μορφή ενός γραμμικού δυναμικού συστήματος [51] ή με μετασχηματισμό επιφάνειας [60].

6.2.3 Διαδικασία προεπεξεργασίας

Η χρήση της χαρακτηριστικής μεταβολής του σχήματος ή του μεγέθους του αντικειμένου αερίου ανίχνευσης πριν την διαδικασία ταξινόμησης μπορεί να αποβεί ιδιαίτερα χρήσιμη. Τα αέρια παρουσιάζουν μια χαοτική δυναμική κίνηση που δεν ακολουθεί κάποιο πρότυπο, ενώ μπορεί να παρουσιάσει μεγάλες διακυμάνσεις ανάμεσα σε λίγα στιγμιότυπα ανάλογα με τις αντίστοιχες συνθήκες του επικρατούν στο εξεταζόμενο σενάριο [53]. Μία τέτοια εφαρμογή απαιτεί ένα αρκετά ισχυρό σύστημα ανίχνευσης κίνησης, που θα μπορούσε να χαρακτηρίζει, να ομαδοποιεί και να παρακολουθεί τα κινούμενα αντικείμενα, προκειμένου να εξαλειφθεί το πρόβλημα των στατικών προηγούμενα κινούμενων αντικειμένων που "λιώνουν" στο παρασκήνιο, όπως με τη χρήση του αλγορίθμου αφαίρεσης υποβάθρου μέσω ενός επιλεκτικά μεταβαλλόμενου μοντέλου. Επιπλέον, μία κανονικοποίηση της αντίθεσης των στιγμιότυπων του βίντεο, ή της εφαρμογής διόρθωσης γάμα (gamma correction) μεταξύ των εικόνων κατά το στάδιο της προεπεξεργασίας θα μπορούσε να είναι επωφελής. Πιο εξελιγμένοι κανόνες εξάλειψης, κατά το στάδιο της μετά-επεξεργασίας που λαμβάνουν υπόψη τα αποτελέσματα της ανίχνευσης στη χωρική αλλά και χρονική γειτονιά του εκάστοτε κελιού, θα μπορούσε να συμβάλει θετικά στην μείωση των εσφαλμένων θετικών ανιχνεύσεων [61].

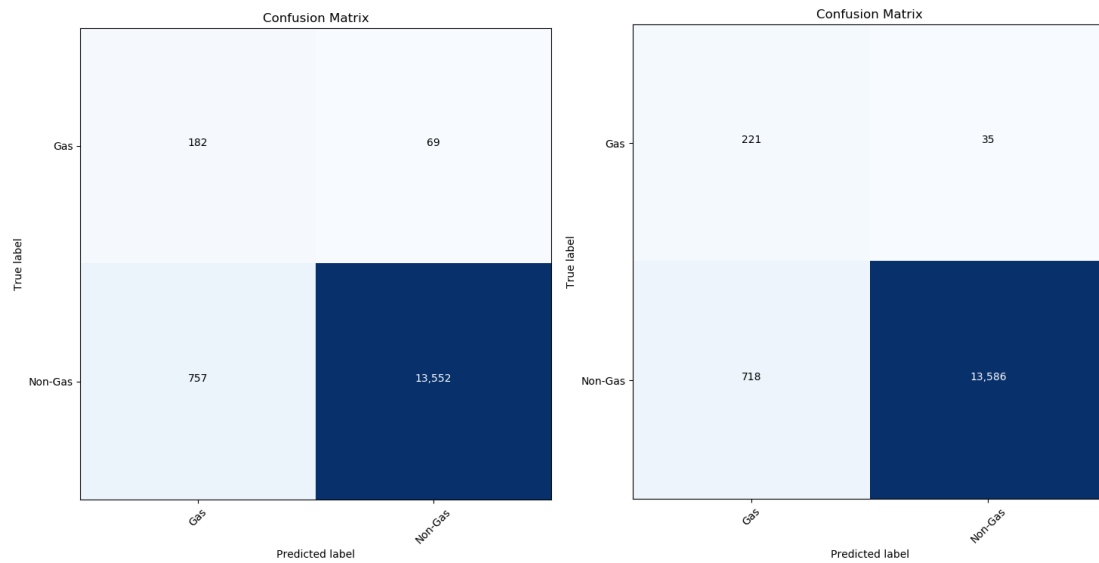
6.2.4 Κινούμενη θερμική κάμερα

Τέλος, η παρούσα εργασία θα μπορούσε να αποτελέσει τη βάση για την εξέλιξη των αλγοριθμικών διαδικασιών ανίχνευσης με κινούμενη θερμική κάμερα. Σχετικά πρόσφατα οι Yusuf Hakan Habiboglu, Osman Günay και A. Enis Çetin δημοσίευσαν μία τεχνική ανίχνευσης αντικειμένων μεταβαλλόμενου σχήματος, με τη χρήση μίας αργά κινούμενης κάμερας [61] και ενός συνδυασμού περιγραφών υφής και χρώματος. Καθώς παρουσιάζεται στην παρούσα έρευνα η χρήση περιγραφών χαρακτηριστικών κίνησης, θα μπορούσε σε έναν βαθμό να προσαρμοστεί σε ένα τέτοιο σενάριο. Μία κίνηση της κάμερας μπορεί να αντισταθμιστεί μη με τις κατάλληλες αλγοριθμικές διαδικασίες σταθεροποίησης εικόνας, ενώ ο αλγόριθμος μπορεί να σταματά στις απότομες μεταβολές της ταχύτητας της κίνησης της κάμερας, καθώς η διαδικασία ανίχνευσης δεν θα είχε χρηστικότητα σε τέτοιες περιπτώσεις.

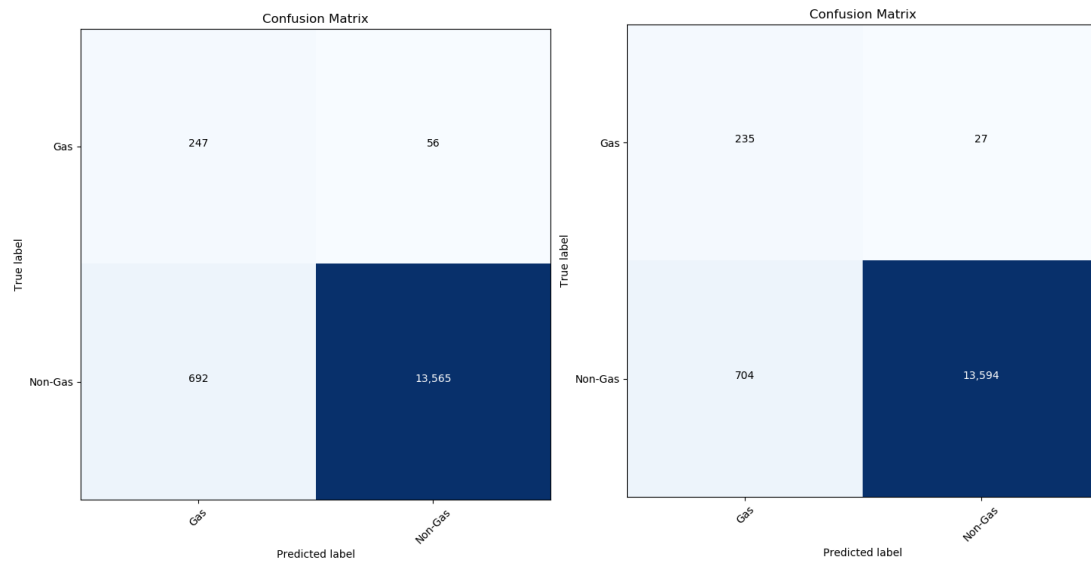
6.2.5 Βελτιώσεις στη μέθοδο εξαγωγής υποβάθρου

Η τεχνική αυτή βασίζεται κυρίως στο συνδυασμό δύο βίντεο διαφορετικών φασματικών ζωνών, ενός στο μέσο υπέρυθρο και ενός στο ορατό φάσμα. Η τεχνική αυτή απαιτούσε τον συγχρονισμό των βίντεο τόσο σε ρυθμό στιγμιότυπων ανά δευτερόλεπτο (fps), όσο και σε εύρεση του αντίστοιχου στιγμιότυπου από το ένα βίντεο στο δεύτερο, καθώς η λήψη των δύο βίντεο δεν ήταν ενιαία, καθώς υπάρχουν συγκεκριμένα διαστήματα που οι δύο κάμερες κατέγραψαν την ίδια σκηνή. Εκτός από τις χρονικές διαφοροποιήσεις, οι δύο κάμερες λάμβαναν εικόνες διαφορετικού μεγέθους, λόγω διαφορετικών προδιαγραφών, και διαφορετικής γωνίας καθώς είχαν τοποθετηθεί με ένα διάστημα ορισμένων εκατοστών μεταξύ τους. Αυτοί οι παράγοντες θα μπορούσαν να ληφθούν υπόψη πριν την εκτέλεση του πειράματος σε περίπτωση μελλοντικής λήψης νέου συνόλου δεδομένων και να προγραμματιστούν, στο βαθμό που αυτό είναι εφικτό, εκ των προτέρων, πριν την καταγραφή μιας νέας σκηνής, με σκοπό την αποφυγή του υψηλού χρόνου επεξεργασίας των δύο βίντεο κατά τη διαδικασία του συγχρονισμού. Επιπρόσθετα, προτείνεται οι δύο κάμερες να βρίσκονται όσο το δυνατόν πιο κοντά για τη μείωση της γωνιακής διαφοράς της σκηνής. Στο κομμάτι της αλγοριθμικής διαδικασίας που αναπτύχθηκε, θα μπορούσε να γίνει εφαρμογή περισσότερων τεχνικών αφαίρεση υποβάθρου σκηνής και συνεπακόλουθος ο συνδυασμός τους προκειμένου βελτιστοποιηθεί η απομόνωση του αερίου. Τέλος προτείνεται η δοκιμή και εφαρμογή διαφορετικών αλγορίθμων κατάτμησης, εκτός του αλγορίθμου ανάλυσης συνδεδεμένων στοιχείων που εφαρμόζεται στην παρούσα τεχνική, όπως του αλγορίθμου Watershed ή της τεχνικής Region-Growing, που χρησιμοποιούνται σε μια σειρά από εφαρμογές ανάλυσης εικόνας.

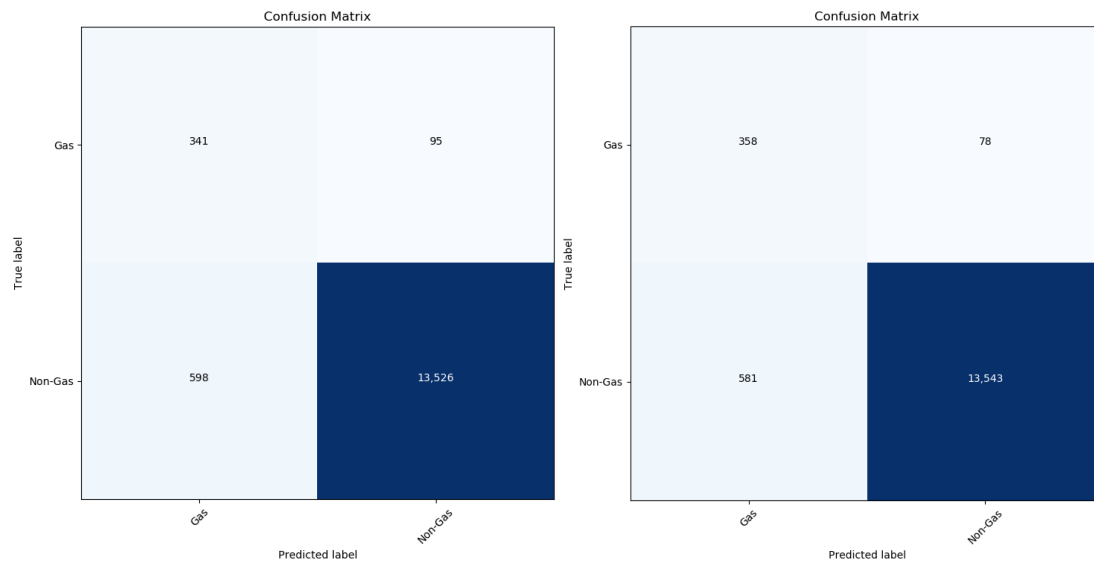
Παράρτημα Α



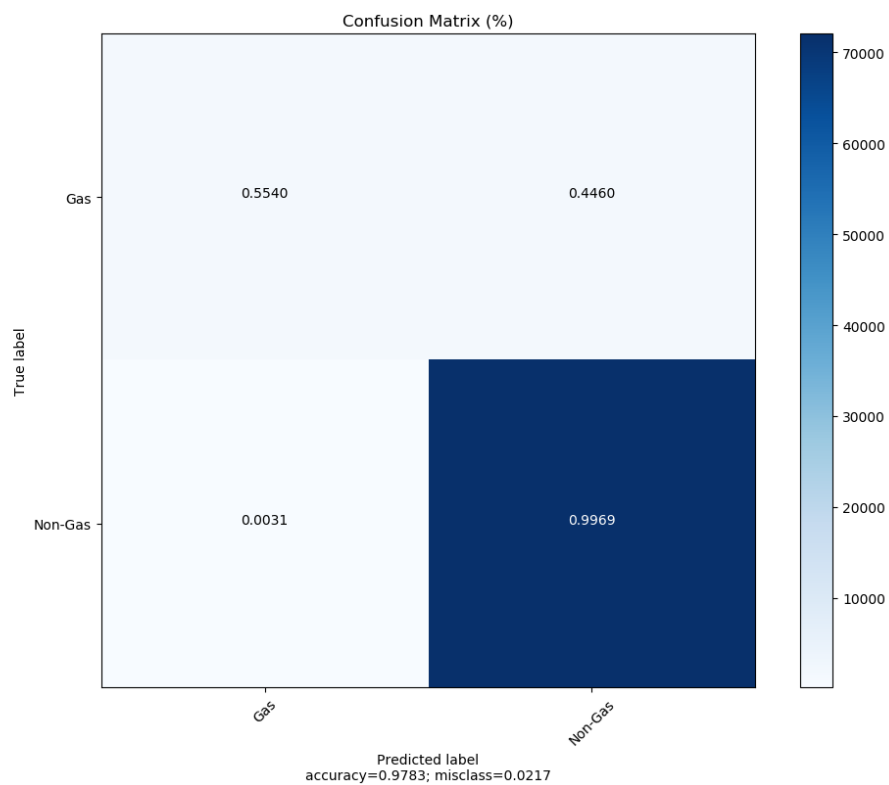
Εικόνα 0.1: Οι παραγόμενοι πίνακες σύγχυσης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών (αριστερά) και δυναμικών (δεξιά) ταξινομητών SVM.



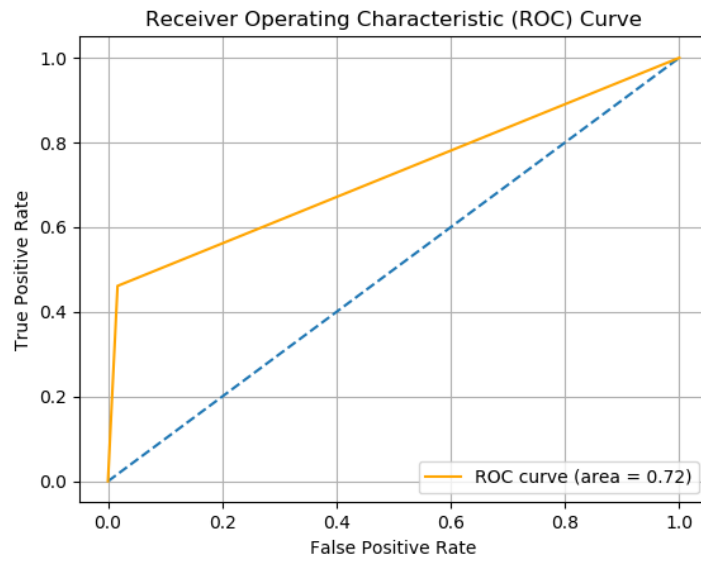
Εικόνα 0.2: Οι παραγόμενοι πίνακες σύγχυσης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών (αριστερά) και δυναμικών (δεξιά) ταξινομητών AdaBoost.



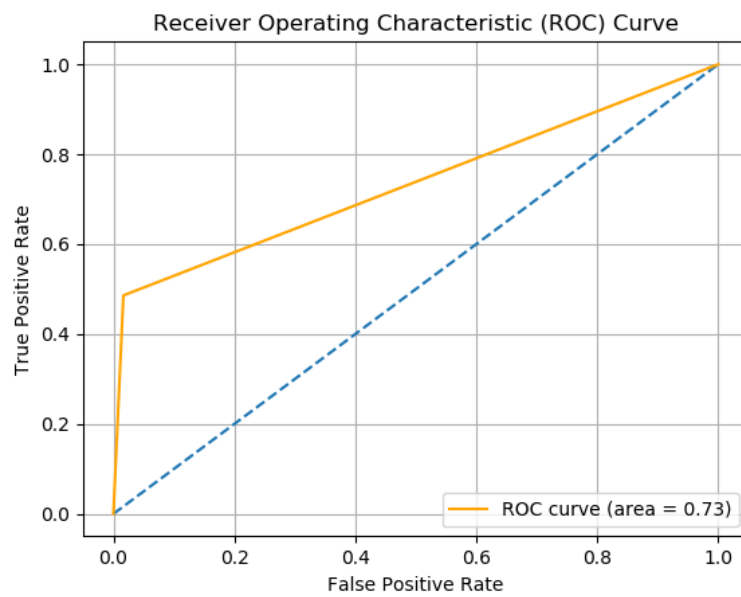
Εικόνα 0.3: Οι παραγόμενοι πίνακες σύγχυσης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών και δυναμικών ταξινομητών **SVM** (αριστερά) και **AdaBoost** (δεξιά).



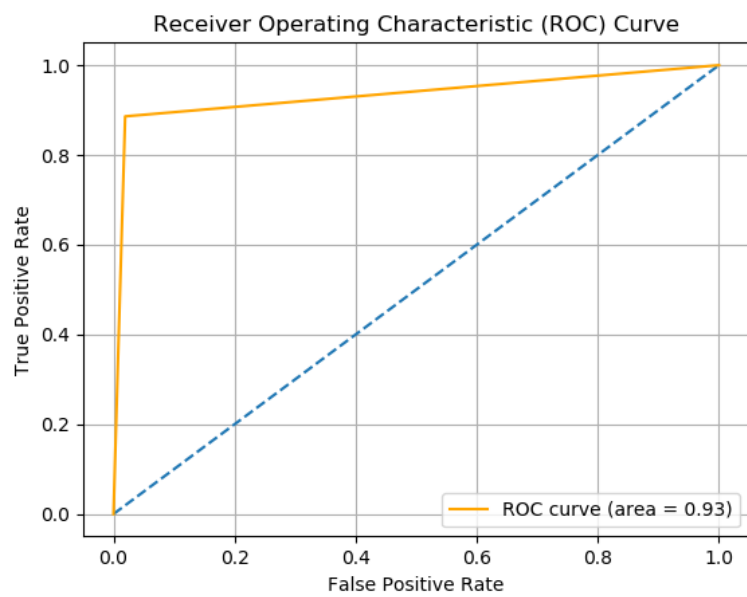
Εικόνα 0.4: Ο παραγόμενος πίνακας σύγχυσης της απόδοσης της μεθόδου ανίχνευσης αερίου με αφαίρεση υποβάθρου.



Εικόνα 0.5: Η καμπύλη ROC της ταξινόμησης του περιγραφέα ιστογράμματος προσανατολισμού (EOH) από SVM.



Εικόνα 0.6: Η καμπύλη ROC της ταξινόμησης από τον συνδυασμό στατικών και δυναμικών ταξινομητών AdaBoost.



Εικόνα 0.7: Η καμπύλη ROC της μεθόδου ανίχνευσης αερίου με αφαίρεση υποβάθρου.





Εικόνα 0.8: Στιγμιότυπα του θερμικού βίντεο απεικόνισης του αέριου προπανίου.

Βιβλιογραφία

- [1] A. R. Krochmal, G. S. Bakken και T. J. LaDuc, «Heat in evolution's kitchen: evolutionary perspectives on the functions and origin of the facial pit of pitvipers,» *Journal of Experimental Biology*, 2004. [1.1]
- [2] N. Arora, D. Martins, D. Ruggerio, E. Tousimis, A. J. Swistel, M. P. Osborne και R. M. Simmons, «Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer,» *The American Journal of Surgery*, p. 523 – 526, 2008. [1.1]
- [3] «sjelectronics,» [Ηλεκτρονικό]. Available: <https://www.sjelectronics.co.uk/blog/cooled-uncooled-thermal-camera>. [Πρόσβαση 20 March 2019]. [1.1]
- [4] M. Vollmer και M. Klaus-Peter , *Infrared Thermal Imaging: Fundamentals, Research and Applications*, 2017. [2.1], [2.3]
- [5] «miniphysics,» [Ηλεκτρονικό]. Available: https://www.miniphysics.com/electromagnetic-spectrum_25.html. [Πρόσβαση 20 April 2019]. [2.1]
- [6] T. E. Avery και G. L. Berlin, *Fundamentals of remote sensing and airphoto interpretation*, New York : Maxwell Macmillan International, 1992. [2.1]
- [7] «Wikipedia - The Free Encyclopedia,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Black-body_radiation. [Πρόσβαση 13 March 2019]. [2.1]
- [8] M. Vollmer και K.-P. Möllmann, «Teaching physics and understanding infrared thermal imaging,» σε *14th Conference on Education and Training in Optics and Photonics: ETOP 2017*, Hangzhou, China, 2017. [2.1]
- [9] «Nasa Earth Observatory,» [Ηλεκτρονικό]. Available: <https://earthobservatory.nasa.gov/features/FalseColor>. [Πρόσβαση 12 March 2019]. [2.2]
- [10] «Opto Engineering. Introduction to infrared vision – Primer on IR theory.,» 2015. [Ηλεκτρονικό]. Available: <https://www.opto-e.com/resources/infrared-theory>. [2.2]
- [11] D. C. Strachan, N. A. Heard, W. J. Hossack, J. F. Boyce και T. M. Cresswell, «Imaging of hydrocarbon vapours and gases by infrared thermography,» *Journal of Physics E: Scientific Instruments*, p. 492, 1985. [2.2]
- [12] «Azooptics,» April 2014. [Ηλεκτρονικό]. Available: <https://www.azooptics.com/Article.aspx?ArticleID=867>. [Πρόσβαση 12 March 2019]. [2.2]
- [13] «Fire and gas systems,» [Ηλεκτρονικό]. Available: http://www.fireandgassystems.com/downloads/manuals/gasimaging/SecondSight_MANUAL.pdf. [2.3]
- [14] M. Mohri, . R. Afshin και T. Ameet , *Foundations of Machine Learning*, The MIT Press, 2012. [3.5]
- [15] K. Bennett και C. Campbell, «Support Vector Machines: Hype or Hallelujah?,» *SIGKDD Explor. Newsl.*, pp. 1-13, 2000. [3.5.1]
- [16] S. Abe, *Support Vector Machines for Pattern Classification (Advances in Pattern Recognition)*, Berlin: Springer-Verlag, 2005, pp. 3.3, 3.3.1. [3.5.1]
- [17] T. Fletcher, «Support vector machines explained,» 23 December 2008. [Ηλεκτρονικό]. Available: https://cling.csd.uwo.ca/cs860/papers/SVM_Explained.pdf. [3.5.1]
- [18] C. J. Burges, «A Tutorial on Support Vector Machines for Pattern Recognition,» *Data Mining and Knowledge Discovery*, pp. 121-167, 1998. [3.5.1]

- [19] Y. Jing, C. Feng και Z. Weidong, «Visual-Based Smoke Detection Using Support Vector Machine,» σε *Natural Computation, 2008. ICNC '08. Fourth International Conference on Natural Computation*, Jinan, China , 2008. [3.5.1]
- [20] J. Shawe-Taylor και N. Cristianini, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods.*, Cambridge: Cambridge University Press, 2000. [3.5.1]
- [21] S. Hua και S. Zhirong, «Support vector machine approach for protein subcellular localization prediction.,» *Bioinformatics*, pp. 721-728, September 2001. [3.5.1]
- [22] P. Viola και M. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features,» σε *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, USA , 2001. [3.5.2]
- [23] K. Ikeuchi, *Computer Vision: A Reference Guide*, Springer, 2014. [3.5.2]
- [24] R. Polikar, «Ensemble based systems in decision making,» *IEEE Circuits and Systems Magazine*, pp. 21-45, October 2006. [3.5.2]
- [25] Z. Wang, «Shape based joint detection and tracking with adaptive multi-motion model and its application in large lump detection,» *University of Alberta Libraries, Alberta*, 2012. [3.5.2], [3.9.2]
- [26] T. G. Dietterich, «An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization,» *Machine Learning*, p. 139–157, August 2000. [3.5.2]
- [27] D. G. Lowe, «Distinctive Image Features from Scale-Invariant Keypoints,» *International Journal of Computer Vision*, pp. 91-110 , November 2004. [3.6.1]
- [28] T. Lindeberg, «Scale-Space Theory: A Basic Tool for Analysing Structures at Different Scales,» *Journal of Applied Statistics*, pp. 224-270, September 1994. [3.6.1]
- [29] . M. A. Fischler και R. C. Bolles, «Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,» *Communications of the ACM*, pp. 381-395 , June 1981. [3.6.2]
- [30] P. Bouchy, A. Goy και P. Andre, «LNG leak detection by infrared thermal camera installation at fos-cavaou terminal,» σε *LNG 17 International Conference & Exhibition on Liquefied Natural Gas Poster Sessions*, Houston, USA, 2013. [3.9.1]
- [31] P. Bernascolle, G. Pelous και L. Grenier, «Remote detection of natural gas clouds in open fields by IR imagery,» *Thermosense XXI*, pp. 409-416, 19 March 1999. [3.9.1]
- [32] N. Edward , S. Baliga και P. Bernascolle, «IR gas imaging in an industrial setting,» *Proc. SPIE 7661*, 3 March 2010. [3.9.1]
- [33] J. Sandsten, H. Edner και S. Svanberg , «Gas visualization of industrial hydrocarbon emissions,» *Optics Express*, pp. 1443-1451, 2004. [3.9.1]
- [34] V. Farley, M. Chamberland, P. Lagueux, A. Vallières, A. Villemaire και J. Giroux, «Chemical agent detection and identification with a hyperspectral imaging infrared sensor,» *Proc. SPIE 6661, Imaging Spectrometry XII*, 12 September 2007. [3.9.1]
- [35] J. Broadwater, T. Spisz και A. Carr, «Detection of gas plumes in cluttered environments using long-wave infrared hyperspectral sensors,» *Proceedings of SPIE - The International Society for Optical Engineering*, April 2008. [3.9.1]
- [36] T. Gerhart, J. Sunu, L. Lieu, E. Merkurjev, J.-M. Chang, J. Gilles και A. Bertozzi, «Detection and tracking of gas plumes in LWIR hyperspectral video sequence data,» *Proc. SPIE 8743, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, 18 May 2013. [3.9.1]
- [37] W. Hirst, «Method and system for rendering visible a plume of dispersing fluid so as to reveal its source». Ευρεσιτεχνία PCT/EP2014/071396, 16 April 2015. [3.9.1]

- [38] S. Soldan, «Sensordatenfusionsansätze in der Thermografie zur Verbesserung der Messergebnisse,» *Dissertation, Schriftenreihe Mess- und Regelungstechnik der Universität Kassel*, 2014. [3.9.1]
- [39] S. Soldan και A. Kroll, «Towards Automated Gas Leak Detection Using IR Gas Imaging Cameras,» σε *AITA 2013 - Advanced Infrared Technology and Applications*, Turin, 2013. [3.9.1]
- [40] H. Abdel-Moati, . J. M. Morris, R. Yingjun και Z. Yousheng , «Remote Gas Detection System Using Infrared Camera Technology and Sophisticated Gas Plume Detection Computer Algorithm,» σε *SPE Annual Technical Conference and Exhibition*, 2013. [3.9.1]
- [41] J. Cheben, Y. Zeng, J. Morris και Y. Ruan, «Differential infrared imager for gas plume detection». US Ευρεσιτεχνία PCT/US2012/028791 , 14 October 2012. [3.9.1]
- [42] D. Manolakis, «Standoff Detection and Identification of Chemical Plumes with Long Wave Hyperspectral Imaging Sensors,» MIT Lincoln Laboratory, Boston, 2009. [3.9.1]
- [43] S. Calderara, P. Piccinini και R. Cucchiara, «Vision based smoke detection system using image energy and color information,» *Machine Vision and Applications*, pp. 705-719, July 2011. [3.9.2]
- [44] E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, H. Habiboğlu, U. Töreyn και S. Verstockt, «Video fire detection – Review,» *Digital Signal Processing*, pp. 1827-1843, December 2013. [3.9.2]
- [45] S. Verstockt, P. Lambert, R. Van de Walle, B. Merci και B. Sette, «State of the art in vision-based fire and smoke detection,» σε *International Conference on Automatic Fire Detection, 14th, Proceedings*, Duisburg, Germany, 2009. [3.9.2]
- [46] C.-Y. Lee, C.-T. Lin, C.-T. Hong και M.-T. Su, «Smoke detection using spatial and temporal analyses,» *International Journal of Innovative Computing, Information and Control*, pp. 4749-4770, July 2012. [3.9.2], [6.2.2]
- [47] J. Chen, Y. You και Q. Peng, «Dynamic Analysis for Video Based Smoke Detection,» *International Journal of Computer Science Issues*, p. 298–304, March 2013. [3.9.2]
- [48] B. Ko, J. Park και J.-Y. Nam, «Spatiotemporal bag-of-features for early wildfire smoke detection,» *Image and Vision Computing*, 786-795 October 2013. [3.9.2]
- [49] F. Yuan, «A fast accumulative motion orientation model based on integral image for video smoke detection,» *Pattern Recognition Letters*, pp. 925-932, May 2008. [3.9.2]
- [50] K. Avgerinakis, I. Kompatsiaris και A. Briassouli, «Smoke detection using temporal HOGHOF descriptors and energy colour statistics from video,» σε *International Workshop on Multi-Sensor Systems and Networks for Fire Detection and Management*, 2012. [3.9.2]
- [51] P. Barmpoutis, K. Dimitropoulos και N. Grammalidis, «Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition,» σε *22nd European Signal Processing Conference (EUSIPCO)*, 2014. [3.9.2], [6.2.2]
- [52] U. Toreyin, Y. Dedeoglu και E. Cetin, «Contour based smoke detection in video using wavelets,» σε *14th European Signal Processing Conference*, Florence, 2006. [3.9.2]
- [53] T. X. Tung και J.-M. Kim, «An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems,» *Fire Safety Journal*, pp. 276-282, July 2011. [3.9.2], [6.2.3]
- [54] J. Gubbi, S. Marusic και M. Palaniswami, «Smoke detection in video using wavelets and support vector machines,» *Fire Safety Journal*, pp. 1110-1115, November 2009. [3.9.2]
- [55] F. Yuan, «Video-based smoke detection with histogram sequence of LBP and LBPV pyramids,» *Fire Safety Journal*, p. 132 – 139, April 2011. [3.9.2], [6.2.2]
- [56] F. Yuan, «A double mapping framework for extraction of shape-invariant features based on multi-scale partitions with adaboost for video smoke detection.,» *Pattern Recognition*, pp. 4326-4336, 01 December 2012. [3.9.2], [5.3]

- [57] N. Dalal, B. Triggs και C. Schmid, «Human Detection Using Oriented Histograms of Flow and Appearance,» σε *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision*, Graz, Austria, 2006. [4.2.3]
- [58] R. Gonzalez-Gonzalez, V. Alarcon-Aquino, R. Rosas-Romero, O. Starostenko, J. Rodriguez-Asomoza και J. M. Ramirez-Cortes, «Wavelet-based smoke detection in outdoor video sequences,» σε *53rd IEEE International Midwest Symposium on Circuits and Systems*, 2010. [6.2.2]
- [59] R. J. Ferrari, H. Zhang και C. Kube, «Real-time detection of steam in video images,» *Pattern Recognition*, pp. 1148-1159 , March 2007. [6.2.2]
- [60] W. Ye, J. Zhao, S. Wang, Y. Wang, . D. Zhang και . Z. Yuan, «Dynamic texture based smoke detection using Surfacelet transform and HMT model,» *Fire Safety Journal*, pp. 91-101, 2015. [6.2.2]
- [61] Y. H. Habiboğlu, O. Günay και A. E. Çetin, «Covariance matrix-based fire and flame detection method in video,» *Machine Vision and Applications*, p. 1103–1113, November 2012. [6.2.3], [6.2.4]

Λίστα Εικόνων

Εικόνα 1.1: Θερμική απεικόνιση διαρροής αναισθητικού αερίου από μια μάσκα [3].2
Εικόνα 1.2: Σημειακός αισθητήρας ανίχνευσης μεθανίου (CH_4) και προπανίου (C_3H_8) της εταιρείας EURO-GAS.3
Εικόνα 1.3: Ενισχυμένη θερμική απεικόνιση υψηλής ευαισθησίας με χρήση θερμικής κάμερας Ventus της εταιρείας Sierra Olympic.4
Εικόνα 2.1: Το ηλεκτρομαγνητικό φάσμα και οι επιμέρους κατηγορίες του [5].6
Εικόνα 2.2: Διαρροή προπανίου που λήφθηκε με τη θερμική κάμερα Ventus της Sierra Olympic.7
Εικόνα 2.3: Η ακτινοβολία του αντικειμένου (κόκκινη γραμμή), η ακτινοβολία της ατμόσφαιρας (μπλε γραμμή) και η ανάκλαση της ακτινοβολίας από τον περιβάλλοντα χώρο (πράσινη γραμμή) [4].7
Εικόνα 2.4: Κατανομή της ακτινοβολίας μαύρου σώματος για διαφορετικές θερμοκρασίες σύμφωνα με τον νόμο του Planck [7].8
Εικόνα 2.5: Εκπομπή (αριστερά) και η θερμική ακτινοβολία (δεξιά) ως συνάρτηση του μήκους κύματος ενός μελανού σώματος, ενός γκρι σώματος και ενός επιλεκτικού πομπού [8].9
Εικόνα 2.6: Τα ατμοσφαιρικά παράθυρα στο ηλεκτρομαγνητικό φάσμα [9].10
Εικόνα 2.7: Σχετικές φασματικές αποκρίσεις ανιχνευτών φωτονίων και θερμικών ανιχνευτών [12].11
Εικόνα 2.8: Το ποσοστό μετάδοσης ακτινοβολίας των αερίων του διοξειδίου του άνθρακα, του μεθανίου, του βενζολίου και του εξαφθοριούχου θείου αντίστοιχα [4].12
Εικόνα 2.9: Ανίχνευση ψυχρών αερίων μπροστά από ένα θερμό υπόβαθρο με ή χωρίς φίλτρο λόγω της απορρόφησης. Οι περιοχές με την κόκκινη και πράσινη σκιά αντικατοπτρίζουν την ανιχνευθείσα φασματική περιοχή του συστήματος της κάμερας [4].13
Εικόνα 2.10: Αύξηση σήματος λόγω της διάχυσης της υπέρυθρης ακτινοβολίας από ένα θερμό αντικείμενο στα μόρια του αερίου [4].14
Εικόνα 2.11: Σύστημα θερμικής κάμερας ανίχνευσης αερίων [13].14
Εικόνα 3.1 Κυκλικές συμμετρικές γειτονιές για διαφορετικές τιμές P και R ((8,1),(8,2) και (16,2) αντίστοιχα).20
Εικόνα 3.2 Παράδειγμα υπολογισμού περιγραφέα $LBP_{8,1}$ με την αρχική εικόνα (αριστερά), το αποτέλεσμα του κατωφλίου (μέση) με μοτίβο=11110001 και τα βάρη από το κωδικοποιημένο άθροισμα (δεξιά), με αποτέλεσμα $LBP = 1 + 16 + 32 + 64 + 128 = 241$20
Εικόνα 3.3: Κατασκευή ιστογράμματος (8,1) LBP, όπου η τιμή του τελεστή LBP κάθε εικονοστοιχείου υπολογίζεται (αριστερά) και απεικονίζεται στο ιστόγραμμα (δεξιά).21

Εικόνα 3.4: Ομοιόμορφα ανεξάρτητα περιστροφής πρότυπα γειτονιάς (8,P) όπου οι μαύρες βούλες αντιστοιχούν σε 0 και οι άσπρες σε 1.	21
Εικόνα 3.5: Αποσύνθεση ενός επιπέδου διακριτού σήματος από μία εικόνα με εφαρμογή φίλτρων. Τα L_0 και H_1 υποδηλώνουν τα φίλτρα υψηλής και χαμηλής διαπερατότητας ενώ τα $2 \downarrow 1$ και $1 \downarrow 2$ υποδηλώνουν την υποδειγματοληψία των στηλών και των γραμμών της εικόνας αντίστοιχα.	22
Εικόνα 3.6: Αποσύνθεση τριών επιπέδων με χρήση κυματιδίων (Wavelets) μιας εικόνας.	23
Εικόνα 3.7: Διαδικασία κατασκευής ιστογράμματος του περιγραφέα βάθμωσης.	24
Εικόνα 3.8: Η συνένωση ιστογραμμάτων επικαλυπτόμενων block.	25
Εικόνα 3.9: Διαδικασία υπολογισμού περιγραφέα οπτικής ροής.	26
Εικόνα 3.10: Το διάνυσμα οπτικής ροής ενός κινούμενου αντικειμένου σε μια ακολουθία στιγμιότυπων βίντεο.	26
Εικόνα 3.11: Ο περιγραφέας τοπικών δυαδικών προτύπων τριών διαστάσεων.	27
Εικόνα 3.12: Γραμμικά διαχωρίσιμα (αριστερά) και μη γραμμικά διαχωρίσιμα (δεξιά) δεδομένα δυσδιάστατου χώρου, συμπεριλαμβανομένων των συναρτήσεων απόφασης αντίστοιχα.	29
Εικόνα 3.13: Οπτικοποίηση ταξινόμησης με διανύσματα υποστήριξης ενός γραμμικά διαχωρίσιμου δυαδικού συνόλου. Τα τετράγωνα και οι κύκλοι στο γράφημα αντιπροσωπεύουν τις δύο κλάσεις +1 και -1 αντίστοιχα.	31
Εικόνα 3.14: Αδιάσπαστο γραμμικά σύνολο δεδομένων με απεικόνιση των "χαλαρών" μεταβλητών ξ_i και ξ_j , των οποίων το μέγεθος καθορίζεται από την αντίστοιχη απόστασή τους από τα όρια της κάθε κλάσης.	33
Εικόνα 3.15: Μη γραμμικά διαχωρίσιμα δεδομένα στο δισδιάστατο χώρο εισόδου (αριστερά) και η προβολή τους στον τρισδιάστατο χώρο χαρακτηριστικών με χρήση συνάρτησης πυρήνα (δεξιά).	34
Εικόνα 4.1: Η πειραματική διάταξη των καμερών.	43
Εικόνα 4.2: Η πειραματική διάταξη (αριστερά) και τα αέρια Μεθάνιο CH_4 και προπάνιο C_3H_8 αντίστοιχα (δεξιά).	44
Εικόνα 4.3: Παράδειγμα στιγμιότυπων raw δεδομένων (αριστερά) και δεδομένων με ενίσχυση ακμών (δεξιά) του αερίου προπάνιου.	44
Εικόνα 4.4: Παράδειγμα στιγμιότυπων raw δεδομένων (αριστερά) και δεδομένων με ενίσχυση ακμών (δεξιά) του αερίου μεθανίου.	45
Εικόνα 4.5: Μεθοδολογία ανίχνευσης αερίου με χρήση ταξινομητών.	46
Εικόνα 4.6: Ενδεικτικά patches εκπαίδευσης με αέριο (αριστερά) και χωρίς αέριο (δεξιά) για την εξαγωγή στατικών περιγραφέων χαρακτηριστικών.	47
Εικόνα 4.7: Παραδείγματα ιστογραμμάτων των τελεστών LBP^{u2} (πάνω) και των LBP^{riu2} (μέση) για τα αντίστοιχα patch αερίου(κάτω αριστερά) και μη αερίου (κάτω δεξιά).	49

Εικόνα 4.8: Ιστογράμματα προσανατολισμού ακμών αερίου (αριστερά) και μη αερίου (δεξιά) με βήμα $K=18$ και ανά 10° .	49
Εικόνα 4.9: Ιστογράμματα οπτικής ροής patch αερίου (αριστερά) και patch μη-αερίου (δεξιά).	50
Εικόνα 4.10: Μεθοδολογία ανίχνευσης αερίου με συνδυασμό μεθόδων αφαίρεσης υποβάθρου.	52
Εικόνα 4.11: Εύρεση ομόλογων σημείων με τον αλγόριθμο SIFT στις δύο εικόνες.	53
Εικόνα 4.12: Η επικαλυπτόμενη περιοχή των δύο εικόνων.	53
Εικόνα 4.13: Φιλτράρισμα αντιστοιχιών με τον αλγόριθμο RANSAC και αφινικός μετασχηματισμός της εικόνας.	54
Εικόνα 4.14: Το στιγμιότυπο του θερμικού με ύπαρξη αερίου (αριστερά, με κόκκινο η περιοχή ύπαρξης αερίου) και το εξαγόμενο υπόβαθρο με τη τεχνική GMM.	55
Εικόνα 4.15: Το στιγμιότυπο του θερμικού με ύπαρξη αερίου (αριστερά, με κόκκινο η περιοχή ύπαρξης αερίου) και το εξαγόμενο υπόβαθρο με τη τεχνική Selective model.	55
Εικόνα 4.16: Το αποτέλεσμα του συνδυασμού των τριών βίντεο υποβάθρου μετά από μορφολογικές διαδικασίες (δεξιά) και το αντίστοιχο στιγμιότυπο του θερμικού βίντεο (αριστερά).	56
Εικόνα 4.17: Αποτέλεσμα ανίχνευσης μετά τις μορφολογικές διαδικασίες και ανάλυσης συνδεδεμένων στοιχείων του βίντεο διαφοράς υποβάθρων (δεξιά, με κόκκινο περίγραμμα τα υπολειπόμενα σχήμα εικονοστοιχείων μη-αερίου) και το αντίστοιχο στιγμιότυπο του θερμικού βίντεο (αριστερά).	56
Εικόνα 4.18: Παράδειγμα γραφικής παράστασης καμπύλης ROC.	58
Εικόνα 5.1: Στιγμιότυπο από το θερμικό βίντεο (αριστερά) και το αποτέλεσμα της ψηφιοποίησης του αερίου για την παραγωγή δεδομένων επικύρωσης (δεξιά).	60
Εικόνα 5.2: Το αποτέλεσμα ψηφιοποίησης (αριστερά) και το αποτέλεσμα της επεξεργασίας σε θετικά πραγματικά κελιά (δεξιά).	61
Εικόνα 5.3: Επιτυχημένες ανιχνεύσεις με χρήση AdaBoost για την ταξινόμηση LBP^{ru2} χαρακτηριστικών.	64
Εικόνα 5.4: Επιτυχημένες ανιχνεύσεις με χρήση SVM για την ταξινόμηση ΕΟΗ χαρακτηριστικών (αριστερά) και συνδυασμού στατικών και δυναμικών SVM ταξινομητών (δεξιά).	64
Εικόνα 5.5: Επιτυχημένες ανιχνεύσεις με χρήση συνδυασμού στατικών και δυναμικών AdaBoost ταξινομητών (αριστερά) και συνδυασμού μόνο δυναμικών AdaBoost ταξινομητών (δεξιά).	65
Εικόνα 5.6: Εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού ταξινομητών SVM (αριστερά) και συνδυασμού ταξινομητών AdaBoost (δεξιά).	65
Εικόνα 5.7: Επιτυχείς και εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού στατικών ταξινομητών AdaBoost (αριστερά) και SVM (δεξιά).	66

Εικόνα 5.8: Επιτυχείς και εσφαλμένες ανιχνεύσεις με χρήση συνδυασμού δυναμικών ταξινομητών AdaBoost (αριστερά) και SVM (δεξιά).....	66
Εικόνα 5.9: Επιτυχείς ανιχνεύσεις εικονοστοιχείων αερίου με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).	67
Εικόνα 5.10: Εσφαλμένες ανιχνεύσεις εικονοστοιχείων αερίου (με κόκκινο περίγραμμα) με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).	68
Εικόνα 5.11: Επιτυχείς και εσφαλμένες (με κόκκινο περίγραμμα) ανιχνεύσεις εικονοστοιχείων αερίου με τη μέθοδο εξαγωγής υποβάθρου (αριστερά) και τα αντίστοιχα στιγμιότυπα (δεξιά).	68
Εικόνα 7.1: Οι παραγόμενοι πίνακες σύγκρισης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών (αριστερά) και δυναμικών (δεξιά) ταξινομητών SVM.	74
Εικόνα 7.2: Οι παραγόμενοι πίνακες σύγκρισης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών (αριστερά) και δυναμικών (δεξιά) ταξινομητών AdaBoost.....	74
Εικόνα 7.3: Οι παραγόμενοι πίνακες σύγκρισης της απόδοσης της ταξινόμησης από τον συνδυασμό στατικών και δυναμικών ταξινομητών SVM (αριστερά) και AdaBoost (δεξιά).	75
Εικόνα 7.4: Ο παραγόμενος πίνακας σύγκρισης της απόδοσης της μεθόδου ανίχνευσης αερίου με αφαίρεση υποβάθρου.....	75
Εικόνα 7.5: Η καμπύλη ROC της ταξινόμησης του περιγραφέα ιστογράμματος προσανατολισμού (EOH) από SVM.....	76
Εικόνα 7.6: Η καμπύλη ROC της ταξινόμησης από τον συνδυασμό στατικών και δυναμικών ταξινομητών AdaBoost.	76
Εικόνα 7.7: Η καμπύλη ROC της μεθόδου ανίχνευσης αερίου με αφαίρεση υποβάθρου.	77
Εικόνα 7.8: Στιγμιότυπα του θερμικού βίντεο απεικόνισης του αερίου προπανίου.	78

Λίστα Πινάκων

Πίνακας 4-1: Πίνακας Σύγχυσης.	57
Πίνακας 5-1: Αποτελέσματα απόδοσης του ταξινομητή SVM για διαφορετικούς περιγραφείς χαρακτηριστικών.....	62
Πίνακας 5-2: Αποτελέσματα απόδοσης του ταξινομητή AdaBoost για διαφορετικούς περιγραφείς χαρακτηριστικών.	62
Πίνακας 5-3: Αποτελέσματα απόδοσης ταξινόμησης από τον συνδυασμό εκπαιδευόμενων ταξινομητών διαφορετικών περιγραφέων χαρακτηριστικών.....	63
Πίνακας 5-4: Αποτελέσματα απόδοσης της κατηγοριοποίησης των εικόνων από τη μέθοδο εξαγωγής υποβάθρου.	67